República Federativa do Brasil
Ministério do Desenvolvimento, Indústria e do Comércio Exterior
Instituto Nacional da Propriedade Industrial

(21) **PI 0810703-3 A2**

BRPI0810703A2

(22) Data de Depósito: 02/05/2008
(43) Data da Publicação: 26/12/2012
(RPI 2190)

(51) Int.Cl.:
C12N 15/82
C07K 14/415

(54) Título: MÉTODO PARA INTENSIFICAR TRAÇOS RELACIONADOS COM RENDIMENTO EM PLANTAS COM RELAÇÃO ÁS PLANTAS DE CONTROLE, MOLÉCULA DE ÁCIDO NUCLEICO ISOLADA, POLIPEPTÍDEO ISOLADO, MÉTODO PARA A PRODUÇÃO DE UMA PLANTA TRANSGÊNICA, PLANTA OU PARTE DA MESMA, CONSTRUÇÃO, E, USO DE UM POLIPEPTÍDEO OU DE UM ÁCIDO NUCLEICO

(30) Prioridade Unionista: 03/05/2007 EP 07107448.8, 29/05/2007 EP 07109052.6, 29/05/2007 EP 07109068.2, 11/06/2007 EP 07109961.8, 19/06/2007 EP 07110548.0, 19/06/2007 EP 07110557.1, 08/08/2007 US 60/916575, 29/06/2007 US 60/937989, 06/06/2007 US 60/942214, 05/07/2007 US 60/948036, 29/05/2007 EP 07109052.6, 29/05/2007 EP 07109068.2, 11/06/2007 EP 07109961.8, 19/06/2007 EP 07110548.0, 19/06/2007 EP 07110557.1, 08/08/2007 US 60/916575, 06/06/2007 US 60/942214, 05/07/2007 US 60/948036, 29/05/2007 EP 07109052.6, 29/05/2007 EP 07109068.2, 11/06/2007 EP 07109961.8, 08/08/2007 US 60/916575, 29/06/2007 US 60/937989, 06/06/2007 US 60/942214, 05/07/2007 US 60/948036, 29/05/2007 EP 07109068.2, 19/06/2007 EP 07110548.0, 19/06/2007 EP 07110557.1, 29/06/2007 US 60/937989, 06/06/2007 US 60/942214, 05/07/2007 US 60/948036, 29/05/2007 EP 07109052.6, 29/05/2007 EP 07109068.2, 11/06/2007 EP 07109961.8, 19/06/2007 EP 07110548.0, 19/06/2007 EP 07110557.1, 29/06/2007 US 60/937989, 05/07/2007 US 60/948036, 11/06/2007 EP 07109961.8, 19/06/2007 EP 07110548.0, 08/08/2007 US 60/916575, 29/06/2007 US 60/937989, 06/06/2007 US 60/942214, 05/07/2007 US 60/948036, 11/06/2007 EP 07109961.8, 08/08/2007 US 60/916575, 06/06/2007 US 60/942214, 05/07/2007 US 60/948036

(73) Titular(es): Basf Plant Science GmbH

(72) Inventor(es): Amber Shirley, Ana Isabel Sanz Molinero, Bryan McKersie, Lalitree Darnielle, Steven Vandenabeele, Valerie Frankard, Yves Hatzfeld

(74) Procurador(es): Momsen, Leonardos & Cia.

(86) Pedido Internacional: PCT US2008005732 de 02/05/2008

(87) Publicação Internacional: WO 2008/137108de 13/11/2008

(57) Resumo: MÉTODO PARA INTENSIFICAR TRAÇOS RELACIONADOS COM O RENDIMENTO EM PLANTAS COM RELAÇÃO ÀS PLANTAS DE CONTROLE, MOLÉCULA DE ÁCIDO NUCLEICO ISOLADA, POLIPEPTÍDEO ISOLADO, MÉTODO PARA A PRODUÇÃO DE UMA PLANTA TRANSGÊNICA, PLANTA OU PARTE DA MESMA, CONSTRUÇÃO, E, USO DE UM POLIPEPTÍDEO OU DE UM ÁCIDO NUCLEICO. A presente invenção diz respeito a, no geral, ao campo da biologia molecular e diz respeito a um método para intensificar traços relacionados com o rendimento e/ou melhorar várias características do desenvolvimento da planta modulando-se a expressão em uma planta de um ácido nucleico que codifica um GRP (Proteína Reguladora do Desenvolvimento). O GRP é selecionado de um domínio LOB que compreende proteína (LOB: Fronteiras de Órgão Lateral), abreviadas aqui como polipeptídeo LBD, um polipetídeo JMJC (JUMONJI-C), um polipeptídeo CKI (Caseína quinase 1), uma proteína semelhante à bHLHí 1 (Hélice-Laço-Hélice básico 11), um polipeptídeo de prolongamento- homeodomínio de homeodomínio de planta (PHDF-HD), um polipeptídeo ASR (abscísico induzido por ácido, tensão e maturação) e/ou um polipeptídeo de fator de transcrição semelhante à proteína de ligação do promotor Squamosa 11 (SPL 11). A presente invenção também diz respeito à plantas tendo expressão modulada de um ácido nucleico que codifica um GRP, cujas plantas tem características de desenvolvimento melhoradas com relação às plantas do tipo selvagem ou outras plantas de controle. A invenção também fornece novos ácidos nucleicos GRP e polipeptídeos GRP bem como construções úteis nos métodos da invenção.

"MÉTODO PARA INTENSIFICAR TRAÇOS RELACIONADOS COM O RENDIMENTO EM PLANTAS COM RELAÇÃO ÀS PLANTAS DE CONTROLE, MOLÉCULA DE ÁCIDO NUCLEICO ISOLADA, POLIPEPTÍDEO ISOLADO, MÉTODO PARA A PRODUÇÃO DE UMA PLANTA TRANSGÊNICA, PLANTA OU PARTE DA MESMA, CONSTRUÇÃO, E, USO DE UM POLIPEPTÍDEO OU DE UM ÁCIDO NUCLEICO"

A presente invenção diz respeito, no geral, ao campo da biologia molecular e diz respeito a um método para intensificar traços relacionados com o rendimento e/ou fornecer a melhora de várias características do desenvolvimento de plantas modulando-se a expressão em uma planta de um ácido nucleico que codifica um GRP (Proteína Reguladora do Desenvolvimento). O GRP é selecionado de um domínio LOB que compreende proteína (LOB: Fronteiras de Órgão Lateral), aqui abreviado como polipeptídeo LBD, um polipeptídeo JMJC (JUMONJI-C), um polipeptídeo CKI (Caseína quinase I), uma proteína semelhante à bHLH11 (Hélice-Laço-Hélice 11 básico), um polipeptídeo de prolongamento-homeodomínio de homeodomínio de planta (PHDf-HD), um polipeptídeo ASR (abscísico induzido por ácido, tensão e maturação) e/ou um polipeptídeo de fator de transcrição semelhante à proteína de ligação de promotor Squamosa 11 (SPL11). A presente invenção também diz respeito a plantas tendo expressão modulada de um ácido nucleico que codifica um GRP, cujas plantas têm características de desenvolvimento melhoradas com relação às plantas do tipo selvagem correspondentes ou outras plantas de controle. A invenção também fornece novos ácidos nucleicos GRP e polipeptídeos GRP bem como construções úteis nos métodos da invenção.

A população mundial sempre crescente e o fornecimento decrescente de terra arável disponível para a pesquisa de combustíveis agrícolas com relação a aumentar a eficácia agrícola. Meios convencionais

para as melhoras de lavoura e horticultura utilizam técnicas de criação seletivas para identificar plantas tendo características desejáveis. Entretanto, tais técnicas de criação seletivas têm diversas desvantagens, isto é, aquelas técnicas são tipicamente trabalho intensivo e resulta em plantas que frequentemente contém componentes genéticos heterogêneos que podem não resultar sempre no traço desejado sendo passado de plantas precursoras. Avanços na biologia molecular têm permitido que a humanidade modifique o *germplasm* de animais e plantas. A engenharia genética de plantas vincula o isolamento e a manipulação de material genético (tipicamente na forma de DNA ou RNA) e a introdução subsequente de material genético em uma planta. Tal tecnologia tem a capacidade de liberar lavouras ou plantas tendo vários traços econômicos, agronômicos ou de horticultura melhorados.

Um traço de interesse econômico particular é o rendimento aumentado. O rendimento é normalmente definido como a produção mensurável de valor econômico de uma lavoura. Isto pode ser definido em termos de quantidade e/ou qualidade. O rendimento é diretamente dependente de diversos fatores, por exemplo, o número e o tamanho dos órgãos, arquitetura da planta (por exemplo, o número de galhos), produção de semente, senescência de folha e mais. O desenvolvimento de raiz, absorção de nutrientes, tolerância à tensão e vigor precoce também podem ser fatores importantes na determinação do rendimento. A otimização dos fatores mencionados acima podem, portanto, contribuir com o aumento do rendimento da lavoura.

O rendimento de semente é um traço particularmente importante, visto que as sementes de muitas plantas são importantes para a nutrição humana e animal. As lavouras, tais como milho, arroz, trigo, canola e soja respondem por mais da metade da absorção calórica humana total, através do consumo direto das sementes por si só ou através do consumo de produtos de carne unidos em sementes processadas. Estas também são uma

fonte de açúcar, óleos e muitos tipos de metabólitos usados em processos industriais. As sementes contém um embrião (a fonte de novos brotos e raízes) e um endosperma (a fonte de nutrientes para o desenvolvimento de embrião durante a germinação e durante o desenvolvimento precoce de mudas). O desenvolvimento de uma semente envolve muitos genes e requer a transferência de metabólitos a partir das raízes, folhas e caules no desenvolvimento de semente. O endosperma, em particular, assimila os precursores metabólicos de carboidratos, óleos e proteínas e os sintetiza em macromoléculas de armazenagem para encher o grão.

Um outro traço importante para muitas lavouras é o vigor precoce. A melhora do vigor precoce é um objetivo importante de programas modernos de criação de arroz tanto em cultivares de arroz temperados quanto tropicais. As raízes longas são importantes para a ancoragem em solo apropriado em arroz semeado na água. Quando o arroz é semeado diretamente em campos alagados e quando as plantas devem emergir rapidamente através da água, raízes mais longas estão associadas com o vigor. Quando a semeadura em sulcos é praticada, mesocotilas e coleoptilas mais longas são importantes para a boa emergência de muda. A capacidade de projetar vigor precoce em plantas deve ser de grande importância na agricultura. Por exemplo, o vigor precoce deficiente foi uma limitação para a introdução de híbridos de milho (Zea mays L.) com base em *germplasm* do Cinturão do Milho no Atlântico europeu.

Um outro traço importante é aquele de tolerância à tensão abiótica melhorada. A tensão abiótica é uma causa primária de perdas de lavoura mundiais, redução dos rendimentos médios para a maioria das plantas de cultivo principais por mais do que 50 % (Wang et al., Plant (2003) 218: 1-14). As tensões abióticas podem ser causadas por secura, salinidade, extremos de temperatura, toxicidade química e tensão oxidativa. A capacidade de melhorar a tolerância da planta à tensão abiótica deve ser de vantagem

econômica grande para o fazendeiros do mundo todo e deve permitir o cultivo de lavouras durante condições adversas e em territórios onde o cultivo de lavouras pode, de outra maneira, não ser possível.

Portanto, o rendimento da lavoura pode ser portanto, aumentado pela otimização de um dos fatores mencionados acima.

Dependendo do uso final, a modificação de certos traços de rendimento pode ser favorecido sobre outros. Por exemplo, para aplicações, tais como produção de forragem ou de madeira ou recurso de biocombustível, um aumento nas partes vegetativas de uma planta pode ser desejável e para aplicações, tais como produção de farinha de amido ou óleo, um aumento em parâmetros de semente pode ser particularmente desejável. Mesmo entre os parâmetros de semente, alguns podem ser favorecidos sobre os outros, dependendo da aplicação. Vários mecanismos podem contribuir para aumentar o rendimento de semente, se aquela estiver na forma de tamanho de semente aumentado ou número de semente aumentado.

Um método para aumentar o rendimento (rendimento de semente e/ou biomassa) em plantas pode ocorrer através da modificação dos mecanismos de desenvolvimento inerentes de uma planta, tal como o ciclo celular ou vários caminhos de sinalização envolvidos no desenvolvimento da planta ou em mecanismos de defesa.

Surpreendentemente, foi observado agora que a modular a expressão de um ácido nucleico que codifica um polipeptídeo GRP selecionado de um domínio LOB que compreende proteína (LOB: Fronteiras de Órgão Lateral), aqui abreviado como polipeptídeo LBD, um polipeptídeo JMJC (JUMONJI-C), um polipeptídeo de caseína quinase, um polipeptídeo de prolongamento-homeodomínio de homeodomínio de planta (PHDf-HD), uma proteína semelhante a bHLH11 (básico Hélice-Laço-Hélice 11), um polipeptídeo ASR (abscícico induzido por ácido, tensão e maturação) e/ou um polipeptídeo de fator de transcrição semelhante à proteína de ligação de

promotor Squamosa 11 (SPL11) dá plantas tendo traços intensificados relacionados com o rendimento e/ou várias características de desenvolvimento de planta melhoradas com relação às plantas de controle.

De acordo com uma forma de realização, é fornecido um método para melhorar ou intensificar traços relacionados com o rendimento e/ou melhorar várias características de desenvolvimento de planta de uma planta com relação às plantas de controle, que compreende modular a expressão de um ácido nucleico que codifica um polipeptídeo GRP selecionado de um domínio LOB que compreende proteína (LOB: Fronteiras de Órgão Lateral), aqui abreviado como polipeptídeo LBD, um polipeptídeo JMJC (JUMONJI-C), um polipeptídeo de caseína quinase, um polipeptídeo de prolongamento-homeodomínio de homeodomínio de planta (PHDf-HD), uma proteína semelhante a bHLH11 (básico Hélice-Laço-Hélice 11), um polipeptídeo ASR (abscícico induzido por ácido, tensão e maturação) e/ou um polipeptídeo de fator de transcrição semelhante à proteína de ligação de promotor Squamosa 11 (SPL11) em uma planta.

FUNDAMENTOS

I. Proteína que compreende o domínio LOB (LOB: Fronteiras de Órgão Lateral)

As proteínas LBD (ou proteínas de domínio LOB) todas dividem um domínio conservado na região de terminal N, conhecida como o domínio LOB. As proteínas do domínio LOB (Shuai et al., Plant Physiol. 129, 747-761, 2002) são observadas em várias espécies vegetais e constituem uma família genética grande: *Arabidopsis* é relatada possuir mais do que 40 genes que codificam as proteínas do domínio LOB, pelo menos 35 genes são encontrados no arroz e pelo menos 15 genes no milho. Os polipeptídeos LBD podem ser reguladores de fatores de transcrição (entre os quais os fatores de transcrição KNOX) são postulados desempenhar um papel na ramificação de capelo e da espiga no milho (Bortiri et al., Plant Cell 18, 574-587, 2006),

formação de raízes adventícias (Liu et al., Plant J. 43, 47-56, 2005; Inukai et al., Plant Cell 17, 1387-1396, 2005), proliferação dos gametófitos femininos (Evans et al., Plant Cell 19, 46-62, 2007), padronização proximal-distal em pétalas (Chalfun-Junior et al., Plant Mol. Biol. 57, 559-575, 2005); morfologia de folha e venação (Iwakawa et al., Plant Cell Physiol. 43, 467-478, 2002). Yang et al. (Molecular Phylogenetics and Evolution 39, 248-262, 2006) três classes diferenciadas de proteínas LBD em arroz, com base na classificação de Iwakawa et al. (2005) e Shuai et al. (2002). Modulação de expressão de gene LOB de classe I (super ou sub-regulação de expressão) frequentemente resulta em efeitos pleiotrópicos, que levam à forma de planta anormal e infertilidade. Por exemplo, o US20060218674 divulga um método para aumentar o tamanho do endosperma em uma semente de planta pela expressão de um polipeptídeo LOB de classe I, entretanto, o tamanho do embrião diminuiu proporcionalmente.

Surpreendentemente, foi observado agora que a modular a expressão de um ácido nucleico que codifica um polipeptídeo LBD dá plantas tendo traços intensificados relacionados com o rendimento, em particular, biomassa aumentada e rendimento de semente, com relação às plantas de controle.

De acordo com uma forma de realização, é fornecido um método para melhorar traços relacionados com o rendimento de uma planta com relação às plantas de controle, que compreende modular a expressão de um ácido nucleico que codifica um polipeptídeo LBD em uma planta. Os traços melhorados relacionados com o rendimento compreenderam biomassa aumentada e rendimento aumentado de semente.

II. polipeptídeo JMJC (JUMONJI-C)

O primeiro gene JUMONJI relatado (que significa cruciforme em japonês) foi identificado pelo aprisionamento do gene em camundongos onde este desempenha um papel essencial no desenvolvimento de tecidos

múltiplos. Até hoje, os polipeptídeos JUMONJI constituem uma classe distinta de proteínas encontradas em procariotas bem como em eucariotas incluindo bactérias, fungos e plantas.

A maioria dos membros da família do polipeptídeo JUMONJI é caracterizada pela presença de um domínio JmjC ou JUMONJI-C. é hipotetizado que durante a evolução, as proteínas antigas que compreende apenas o domínio JmjC adquiriram domínios adicionais que aumentam o espectro dos polipeptídeos JMJC encontrados na natureza. De interesse particular são aqueles polipeptídeos JMJC que adquiriram domínios conservados envolvidos no DNA, RNA e ligação de proteína, tais como dedos de zinco, a proteína de dedo RING rica em FY e os domínios F-box, sugerindo que os polipeptídeos da família JUMONJI podem regular a transcrição, função de cromatona e/ou rotação de proteína. Consequentemente, muitas proteínas JUMONJI em animais foram relatadas afetar o desenvolvimento pelo controle da expressão genética e atividade de cromatina. Por exemplo, um gene JUMONJI de camundongo atua para reprimir a ciclina D1 em embriões e sua atividade é requerida para a cardiogênese normal (Toyoda et al. 2003 Dev Cell. 5(1):85-97.).

Muito progresso foi feito o entendimento do modo de ação das proteínas contendo o domínio jmjC. Cruciais para a sua função biológica podem ser algumas das atividades enzimáticas recentemente reveladas em polipeptídeos JMJC, por exemplo, JHDM1, um polipeptídeo JMJC de origem humana (Tsukada, Nature. 2006; 439(7078):811-6) tem atividade de histona dimetilase e atividade de asparaginil hidroxilase foi relatada em FIH (Fator de Inibição HIF-1 alfa), um fator de transcrição envolvido na resposta celular à hipoxia, (Linke et al. (2004) J. Biol. Chem., Vol. 279, 14391-14397).

Os domínios JmjC tipicamente têm estruturas semelhantes àquelas encontradas em algumas metaloenzimas. A análise estrutural recente do domínio JmjC presente na proteína JUMONJI FIH revelou os resíduos de

aminoácido envolvidos na ligação aos cofatores 2-oxoglutarato e ferro Fe (II) (Dann et al; Proc Natl Acad Sci U S A. 2002; 99(24): 15351-15356). O FIH tem o motivo de ligação de ferro HXD/E característico da maioria das 2-oxoglutarato oxigenases. Além disso e consistente com a atividade de hidrolase, a estrutura secundária de FIH é composta de um núcleo de rocambole de filamento beta que circunda o local de ligação de Fe(II). Estas características são conservadas entre as proteínas contendo domínio JmjC (Trewick et al. EMBO Rep. 2005 (4):315-20).

Em plantas, dois polipeptídeos JMJC estão relatadamente envolvidas no controle do tempo de florescimento. Ambos atuam como repressores dos caminhos de florescimento em *Arabidopsis thaliana* (Noh et al. Plant Cell. 2004. 16(10):2601-13). A atividade enzimática para as proteínas vegetais ainda não foi experimentalmente determinada.

Surpreendentemente, foi observado agora que modular a expressão em uma planta de um ácido nucleico que codifica um JUMONJI-C ou polipeptídeo JMJC dá plantas tendo traços intensificados relacionados com o rendimento, em particular, rendimento aumentado com relação às plantas de controle.

De acordo com uma forma de realização, é fornecido um método para melhorar traços relacionados com o rendimento de uma planta com relação às plantas de controle, que compreende modular a expressão de um ácido nucleico que codifica um polipeptídeo JMJC em uma planta. Os traços melhorados relacionados com o rendimento compreende um ou mais de peso de semente total aumentado por planta, peso de semente milhar (1000 sementes), taxa de enchimento de sementes, índice de colheita, vigor precoce e índice de raiz/broto, ambos sob condições de desenvolvimento ótimas e subótimaa, condições de secura brandas.

III. Caseína quinase I

As proteína quinases representam uma superfamília e os

membros desta superfamília catalisam a transferência reversível de um grupo fosfato de ATP às cadeias secundárias de aminoácido de serina, treonina e tirosina em polipeptídeos alvos. Em particular, a família Caseína quinase 1 (CKI) (EC 2.7.11.1) de função de proteína quinases como reguladores de caminhos de transdução de sinal na maioria dos organismos eucarióticos. Na levedura, a CK I está envolvida na regulação de reparo de DNA e progressão do ciclo celular (Hoekstra MF et al., Science 253: Brockman JL et al., Proc. Natl Acad Sci. USA 89: 9454-9458,1992; Dhillon N and Hoekstra MF, EMBO J 13: 2777-2788,1994). As proteínas de Caseína quinase I são proteína quinases do tipo serina/treonina monomérica que contém um domínio de quinase altamente conservado. Os membros desta família têm extensões de terminal N e terminal C divergentes. A região de terminal N é responsável pelo reconhecimento do substrato e da extensão do terminal C é importante para a interação da quinase com os substratos. A extensão de terminal C também é pensada ser importante para mediar a regulação através da auto-fosforilação (Gross and Anderson, 1 998 Cell Signal 10:699-71 1; Craves and Roach, 1 995, J Biol Chem 270:21689-21694).

Em plantas, diversos membros da família de proteína foram clonados e caracterizados bioquimicamente (Klimczak and Cashmore AR, 1993. Biochem. J. 293: 283-288; Liu et at. 2003, Plant Journal. 36, 189-202; Lee et at. 2005 Plant Cell. 17(10): 2817-2831. O genoma de Arabidopsis foi observado conter pelo menos 14 genes semelhantes à caseína quinase I (CKL). Dentro dos domínios de quinase conservados, os polipeptídeos dividiram 89 % de similaridade de sequência no nível de aminoácido. As 14 isoformas de Arabidopsis CKL ainda foram classificadas nos três grupos com base na localização subcelular. O Grupo 1 localizado predominantemente na periferia celular; grupo 2 no grupo de núcleo 3 o citoplasma. Um CKI de *Nicotiana tabacum* foi localizado no plasmodesmata e propõe desempenhar um papel na comunicação célula a célula (Lee et al. 2005). Outros papéis

propostos para o CKI vegetal é a transdução de sinal em resposta aos estímulos ambientais, o desenvolvimento de raiz e a sensibilidade hormonal vegetal (Liu et al.).

Surpreendentemente, foi observado agora que modular a expressão de um ácido nucleico que codifica um polipeptídeo CKI dá plantas tendo traços intensificados relacionados com o rendimento com relação às plantas de controle.

De acordo com uma forma de realização, é fornecido um método para melhorar traços relacionados com o rendimento de uma planta com relação às plantas de controle, que compreende modular a expressão de um ácido nucleico que codifica um polipeptídeo CKI em uma planta. Os traços melhorados relacionados com o rendimento compreenderam um ou mais de biomassa aumentada, vigor de emergência aumentado e rendimento aumentado de semente.

IV. Polipeptídeo de prolongamento-homeodomínio de homeodomínio de planta (PHDf-HD)

Os fatores de transcrição são usualmente definidos como as proteínas que mostram afinidade de ligação de DNA específicas de sequência e que são capazes de ativar e/ou reprimir a transcrição. O genoma de *Arabidopsis thaliana* codifica quanto a pelo menos 1533 reguladores de transcrição, considerando ~5.9 % de seu número total estimado de genes (Riechmann *et al.* (2000) Science 290: 2105-2109). Os bancos de dados de Fatores de Transcrição Arroz (DRTF) é uma coleção de fatores de transcrição conhecidos e preditos de *Oryza sativa L. ssp. indica* e *Oryza sativa L. ssp. Japonica* e correntemente contém 2.025 modelos genéticos de fatores de transcrição putativos (TF) em *indica* e 2.384 em *japonica*, distribuído em 63 famílias (Gao et al. (2006) Bioinformatics 2006, 22(10):1286-7).

Uma destas famílias é a superfamília de fatores de transcrição de homeodomínio (HD) envolvidos em muitos aspectos de processos de

desenvolvimento. Os fatores de transcrição HD são caracterizados pela presença de um homeodomínio (HD), que é um domínio de ligação de DNA de 60 aminoácidos (BD). *Arabidopsis thaliana* e arroz contém aproximadamente 100 fatores de transcrição HD, que ainda podem ser classificados em subfamílias com base na identidade de sequência de aminoácido (Richardt *et al.* (2007) Plant Phys 143(4): 1452-1466). Algumas destas são caracterizadas pela presença de domínios conservados adicionais que facilitam a ligação de DNA e/ou as interações de proteína-proteína.

Um destes domínios é o dedo PHD, denominado prolongamento de homeodomínio de planta (PHDf) devido à sua associação em um mesmo polipeptídeo com um HD de ligação de DNA, que foi originalmente identificado pela identidade de sequência de aminoácido entre um fator de transcrição de homeobox de milho ZmHOX1a (Bellman & Werr (1992) EMBO J 11:3367-3374) e eu ATHAT3.1 relativo à *Arabidopsis* (Schindler et al. (1993) Plant J 4: 137-150). O PHDf é um motivo semelhante ao motivo de dedo de zinco Cys_4-His-Cys_3 capaz de realizar a quelação de dois íons de zinco. Os PHDfs são encontrados em proteínas nucleares e são pensados estarem envolvidos na regulação transcripcional mediado por cromatina (Halbach *et al.* (2000) Nucleic Acid Res 28(18): 3542-3550).

Os fatores de transcrição que combinam um PHDf e um HD são portanto denominados PHDf-HDs (Halbach *et al.* (2000) *supra).* Em plantas, tais PHDf-HDs ainda são caracterizados pela presença de uma leucina per (ZIP) a montante do PHDf. Ambos os domínios (o ZIP e o PHDf) juntos formam uma região de 180 aminoácidos altamente conservada denominada o domínio ZIP/PHDf (Halbach *et al.* (2000) *supra).*

As plantas de tabaco transgênico que, fortemente, superexpressam cada um dos dois PHDf-HDs de polipeptídeo de milho (ZmHOX1 a ou ZmHOX1 b) usando-se um promotor 35S de vírus mosaico de couve-flôr combinado com um intensificador ômega, mostrou mudanças

morfológicas idênticas: redução de tamanho, formação de raiz adventícia e transformações florais homeóticas (Uberlacker *et al.* (1996) Plant Cell 8: 349-362). As plantas transgênicas de arroz e de tabaco super-expressam fortemente o polipeptídeo HAZ1 PHDF-HD de *Oryza sativa* usando-se o promotor 35S do vírus mosaico de couve-flôr não apresentaram desenvolvimento anormal ou mudança fenotípica em comparação com os tipos selvagens (Ito *et al.* (2004) Gene 331: 9-15).

Na patente US 7.196.245, um polipeptídeo de *Arabidopsis thaliana* PHDf-HD (identificado como G416) foi transformado em Arabidopsis e mostrado promover o florescimento precoce nas plantas transgênicas em comparação com as plantas de controle, sem impacto no rendimento de semente.

Surpreendentemente, foi observado agora que modular, preferivelmente aumentar, a expressão de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD dá plantas tendo traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, com relação às plantas de controle.

De acordo com uma forma de realização, é fornecido um método para intensificar traços relacionados com o rendimento, preferivelmente intensificar traços de sementes relacionados com o rendimento, em plantas com relação às plantas de controle, que compreende modular, preferivelmente aumentar, a expressão de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD em uma planta. Os traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, compreende um ou mais de: número aumentado de panícula primárias, rendimento de semente aumentado total por planta, número aumentado de sementes (enchidas), peso de semente milhar aumentado (TKW), índice de colheita aumentado.

V. Proteína semelhante a bHLH11 (Hélice-Laço-Hélice 11 básico)

Os fatores de transcrição são usualmente definidos como proteínas que mostram a ligação de DNA específica de semente e que são capazes de ativar e/ou reprimir a transcrição. A família do fator de transcrição Hélice-Laço-Hélice básica é uma das famílias maiores de fatores de transcrição que foram caracterizadas em *Arabidopsis thaliana* (Toledo-Ortiz et al., Plant Cell 15, 1749-1770, 2003; Bailey et al., Plant Cell 15, 2497-2501, 2003) e em arroz (Li et al Plant Physiol. 141, 1167-1184, 2006). A característica de distinção da família do fator de transcrição bHLH é a presença de um domínio bipartido que consiste de aproximadamente 60 aminoácidos. Este domínio bipartido é compreendido de uma região básica de ligação de DNA, que liga-se a um E-box de hexanucleotídeo de consenso e duas α-hélices separadas por uma região de arco variável, localizado no terminal Cdo domínio básico. As duas α-hélices promovem a dimerização, permitindo a formação de formação de homo e heterodímeros entre os membros de famílias diferentes. Enquanto o domínio bHLH é evolucionariamente conservado, existe pouca similaridade de sequência entre os clados além do domínio. Li et al. (2006) classifica os fatores de transcrição de arroz e Arabidopsis bHLH em 22 subfamílias, com base na sequência dos domínios bHLH.

Pouco é conhecido sobre a função dos polipeptídeos semelhantes à bHLH11 em plantas. Até agora, apenas um polipeptídeo semelhante à bHLH11, OsPTF1 de arroz, foi caracterizdo. O OsPTF1 é relatado estar envolvido na tolerância à inanição por fosfato (Yi et al., Plant Physiol. 138, 2087-2096). As plantas de arroz que super-expressam o gene sob o controle do promotor 35S não mostra qualquer fenotipo diferente em comparação com as plantas de controle quando desenvolvidas sob condições normais, mas sob condições de limitação de fosfato, as plantas tiveram uma absorção de fosfato melhorada. Sob a limitação de fosfato, as plantas

transgênicas mostraram um aumento na biomassa, teor de fosfato, brotação aumentada e rendimento aumentado de semente.

Surpreendentemente, foi observado agora que modular a expressão I de uma planta de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 dá plantas tendo traços intensificados relacionados com o rendimento, em particular, rendimento aumentado com relação às plantas de controle. Estes efeitos foram mostrados sob condições de desenvolvimento onde o fosfato não foi limitante.

De acordo com uma forma de realização, é fornecido um método para melhorar traços relacionados com o rendimento de uma planta com relação às plantas de controle, que compreende modular a expressão de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 em uma planta. Os traços melhorados relacionados com o rendimento compreenderam rendimento aumentado de semente.

VI. Pantas tendo traços relacionados com o rendimento intensificados e um método para fabricar as mesmas

Os polipeptídeos ASR (abscísico induzido por ácido, tensão e maturação) foram primeiro identificados em tomate (lusem et al 1993 Plant Physiol 102: 1353-1354) com proteínas hidrofílicas altamente carregadas pequenas localizadas nos núcleos de células e ligado à cromatina. A família do gene Asr que codifica polipeptídeos ASR é encontrada disseminada em plantas superiores e os homólogos de ASR foram clonados a partir de um grande número de plantas dicotiledôneas e monocotiledôneas (Carrai et at. 2004 Trends Plant Sci 9: 57-59). A maioria dos genes Asr são super-regulados sob condições de tensão ambiental diferentes, durante a maturação de fruto e no tratamento celular com o hormônio ABA. Os polipeptídeos ASR mostram um alto grau de conservação de sequência (Frankel et at. 2006. Gene Pages 74-83). All known Asr genes contain two highly conserved regions. The first region contains a stretch of His residues at the N-terminus,

possessing sequence-specific Zn^{2+}-dependent atividade de ligação de DNA (Kalifa et al., 2004a Biochem J 381: 373-378). A segunda região é uma parte grande da sequência de terminal C, frequentemente contendo um NLS de sinal de localização nuclear (Cakir et al., 2003 Plant Cell 15: 2165-2180). A proteína ASR1 de tomate é uma proteína intrínsecamente não estruturada que, na ligação de íons de zinco, torna-se ordenada (dobrada) e forma dímeros (Goldgur et al. Plant Physiol. 2007 Feb;143(2):617-28)

Um papel putativo de polipeptídeos ASR na regulação de transcrição genética foi sugerido. Relatadamente, experimentos híbridos de uma levedura revelaram que uma uva ASR liga-se ao promotor de um gene transportador de hexose (VvHT1). Consistentemente, um papel de Asr1 no controle de absorção de hexose em órgãos heterotróficos, tais como tubérculos de batata foi sugerido (Frankel et al. Plant Mol Biol. 2007 Mar;63(5):719-30).A atividade de ligação de DNA dependente de zinco de um membro de proteína da família ASR foi relatado (Kalifa et al. 2004 Biochem J. 2004 Jul 15;381(Pt 2):373-8).

Os domínios de ligação de DNA e zinco de proteína ASR1 foram mapeados (Rom et al. 2006. Biochimie. 88(6):621-8; Goldgur et al. 2007. Plant Physiol. Feb;143(2):617-28).

O uso de polipeptídeos ASR para melhorar os traços agronômicos em plantas foi divulgado, como métodos para intensificar a tolerância de planta a tensões abióticas particulares. Por exemplo, super-expressão em Arabidopsis *(Arabidopsis thaliana)* do gene ortólogo ASR1 *LLA23* de lírios *(Lilium longiflorum)* aumenta a tolerância da planta à secura e à salinidade (Yang et al. 2005. Plant Physiol 139: 836-846). Além disso, a patente US 7.154.025 divulga métodos para aumentar a resistência à tensão por déficit de água aumentando-se a quantidade de proteínas ASR em uma planta.

Surpreendentemente, foi observado agora que modular a

expressão em uma planta de um ácido nucleico que codifica um polipeptídeo ASR dá plantas tendo traços intensificados relacionados com o rendimento sob condições de desenvolvimento sem tensão, com relação às plantas de controle. De acordo com uma primeira forma de realização, a presente invenção fornece um método para intensificar traços relacionados com o rendimento sob condições de desenvolvimento sem tensão em plantas com relação às plantas de controle, que compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo ASR.

VII. 11 semelhante à proteína de ligação do promotor Squamosa (SPL11)

Os polipeptídeos do fator de transcrição são usualmente definidos como proteínas que mostram a afinidade de ligação de DNA específica de sequência e que são capazes de ativar e/ou reprimir a transcrição. Os polipeptídeos de fator de transcrição semelhantes à proteína de ligação do promotor Squamosa (SPL) são proteínas estruturalmente diveras que dividem um domínio de ligação de DNA altamente conservado (DBD) de cerca de 80 resíduos de aminoácido de comprimento (Klein *et al.* (1996) Mol Gen Genet 259: 7-16; Cardon *et al.* (1999) Gene 237: 91-104). O local de ligação de sequência de consenso de DNA de fator de transcrição SPL no promotor de genes alvos é 5'-TNCGTACAA-3' onde N representa qualquer base. Dentro do SPL DBD são dez resíduos de cisteína (Cys) ou histidina (His) conservados (ver Figure 28) dos quais oito são resíduos coordenantes de zinco que ligam dois íons de zinco necessários para a formação de estrutura terciária de dedo de zinco específico de SPL (Yamasaki et al. (2004) J Mol Biol 337: 49-63). Uma segunda característica conservada dentro do SPL DBD é um sinal de localização nuclear bipartido. Fora do DBD, um motivoalvo de micro RNA (miRNA), especialmente alvejado pela família miR156 de miRNAs é encontrado na maioria das sequências de ácido nucleico que codificam os polipeptídeos do fator de transcrição SPL (na região codificadoraou o 3' UTR) por todo o reino vegetal (Rhoades *et al.* (2002) Cell

110: 513-520). A expressão do gene SPL que controla miRNAs pós-transcripcionalmente alvejando mRNAs que codificam SPL para a degradação ou pela repressão translacional.

O genoma de Arabidopsis codifica pelo menos quanto aos 1533 reguladores transcripcionais, levando em conta ~5,9 % de seu número total estimado de genes (Riechmann et al. (2000) Science 290: 2105-2109). Os autores relatam 16 polipeptídeos de fator de transcrição SPL em *Arabidopsis thaliana*, compouca similaridade de sequência entre eles (a parte das características mencionadas acima), o tamanho do polipeptídeo SPL deduzido que varia de 131 a 927 aminoácido. Entretanto, os pares de polipeptídeos de fator de transcrição SPL que dividem homologia de sequência mais alta foram detectados dentro da família SPL desta planta (Cardon et al. (1999)).

Os polipeptídeos de fator de transcrição SPL (encontrados apenas em plantas) caracterizados até hoje mostraram-se funcionar no desenvolvimento da planta, em particular, no desenvolvimento de flôr. As plantas transgênicas que super expressam um polipeptídeo de fator de transcrição SPL3 foram relatadas florescer mais cedo (Cardon et al. (1997) Plant J 12: 367-377). No pedido de patente europeu EP1033405, o ácido nucleico e as sequências de polipeptídeo deduzadas do polipeptídeo de fator de transcrição SPL11 são divulgadas.

Surpreendentemente, foi observado agora que modular a expressão I de uma planta de ácido nucleico que codifica um polipeptídeo SPL11 dá plantas tendo traços intensificados relacionados com o rendimento, em particular, rendimento aumentado com relação às plantas de controle.

De acordo com uma forma de realização, é fornecido um método para intensificar traços relacionados com o rendimento de uma planta com relação às plantas de controle, que compreende modular a expressão de um ácido nucleico que codifica um polipeptídeo SPL11 em uma planta. Os

traços intensificados relacionados com o rendimento compreende um ou mais de vigor de emergência aumentado (vigor precoce de muda aumentado), rendimento de semente total (peso de semente), taxa de enchimento de semente (taxa de enchimento de semente), número de sementes enchidas, número de flores (sementes) por panícula, índice de colheita e peso de semente milhar (1000 sementes), tal aumento ocorrendo tanto sob condições ótimas quanto subótimas de desenvolvimento, preferivelmente condições de secura brandas.

De acordo com uma forma de realização, são fornecidos novos ácidos nucleicos SPL11 e polipeptídeos SPL11 bem como construções que compreendem ácidos nucleicos que codifiquem SPL11, úteis na realização dos métodos da invenção

Definições

Polipeptídeo(s)/Proteína (s)

Os termos "polipeptídeo" e "proteína" são usados de maneira intercambeável aqui e referem-se a aminoácidos em uma forma polimérica de qualquer comprimento, ligados juntos por ligações de peptídeo.

Polinucleotídeo(s)/Ácido(s) nucleico(s)/Sequência(s) de ácido nucleico/Sequência(s) de nucleotídeo

Os termos "polinucleotídeo(s)", "sequência(s) de ácido nucleico", " sequência(s) de nucleotídeo", "ácido(s) nucleico(s)", "molécula de ácido nucleico" são usados de maneira intercambeável neste e referem-se a nucleotídeos, ribonucleotídeos ou desóxirribonucleotídeos ou uma combinação de ambos, em uma forma não ramificada polimérica de qualquer comprimento.

Planta(s) de controle

A escolha de plantas de controle adequadas é de um projeto experimental e pode incluir plantas do tipo selvagem correspondentes ou plantas correspondentes sem o gene de interesse. A planta de controle é

tipicamente das mesmas espécies de plantas ou ainda da mesma variedade como a planta a ser avaliada. A planta de controle também pode ser um nulizigoto da planta a ser avaliada. Os Nulizigotos são indivíduos que perdem o transgene por segregação. Uma "planta de controle" como usada neste refere-se não apenas às plantas otais, mas também à partes de planta, incluindo sementes e partes de sementes.

Homólogo(s)

Os "homólogos" e uma proteína abrange peptídeos, oligopeptídeos, polipeptídeos, proteínas e enzimas tendo substituições, anulações e/ou inserções de aminoácido com relação à proteína não modificada em questão e tendo atividade biológica e funcional similar à proteína não modificada a partir da qual estes são derivados.

Uma anulação refere-se à remoção de um ou mais aminoácidos de uma proteína.

Uma inserção refere-se a um ou mais resíduos de aminoácido sendo introduzidos em um local predeterminado em uma proteína. As inserções podem compreender fusões de terminal N e/ou terminal C bem como inserções intra-sequência de aminoácidos simples ou múltiplos. No geral, as inserções dentro da sequência de aminoácido serão menores do que as fusões de terminal N ou C, da ordem de cerca de 1 a 10 resíduos. Os exemplos de proteínas de fusão ou peptídeos de terminal N ou C incluem o domínio de ligação ou o domínio de ativação de um ativador de transcrição como usado no sistema de dois híbridos de levedura, proteínas de revestimento de fago, (histidina)-6-tag, glutationa S-transferase-tag, proteína A, proteína de ligação de maltose, diidrofolato redutase, epítopo Tag-100, epítopo c-myc, FLAG®-epítopo, lacZ, CMP (peptídeo de ligação de calmodulina), epítopo HA, epítopo de proteína C e epítopo VSV.

Uma substituição refere-se à troca de aminoácidos da proteína por outros aminoácidos tendo propriedades similares (tais como

hidrofobicidade similar, hidrofilicidade, antigenicidade, propensão a formar ou quebrar estruturas α-hélicas ou estruturas de 8 lâminas). As substituições de aminoácido são, tipicamente, de resíduos simples, mas podem ser agrupadas dependendo das restrições funcionais colocadas no polipeptídeo; as inserções serão, usualmente da ordem de cerca de 1 a 10 resíduos de aminoácido. As substituições de aminoácido são preferivelmente substituições de aminoácido conservativas. As tabelas de substituição conservativas são bem conhecidas na técnica (ver, por exemplo, Creighton (1984) Proteins. W.H. Freeman and Company (Eds) e Tabela 1 abaixo).

Tabela 1: Exemplos de substituições de aminoácido conservadas

resíduo	substituições conservativas	resíduo	substituições conservativas
Ala	Ser	Leu	Ile; Val
Arg	Lys	Lys	Arg; Gln
Asn	Gln; His	Met	Leu; Ile
Asp	Glu	Phe	Met; Leu; Tyr
Gln	Asn	Ser	Thr; Gly
Cys	Ser	Thr	Ser; Val
Glu	Asp	Trp	Tyr
Gly	Pro	Tyr	Trp; Phe
His	Asn; Gln	Val	Ile; Leu
Ile	Leu, Val		

As substituições, anulações e/ou inserções de aminoácido podem ser feitas facilmente usando-se técnicas sintéticas de peptídeo bem conhecidas na técnica, tais como síntese de peptídeo de fase sólidae outros ou pela manipulação de DNA recombinante. Os métodos para a manipulação de sequências de DNA para a produção de variantes de substituição, inserção ou anulação de uma proteína são bem conhecidos na técnica. Por exemplo, as técnicas para a realização de mutações de substituição em locais predeterminados no DNA são bem conhecidos por aqueles habilitados na técnica e incluem a mutagênese de M13, mutagênese in vitro de T7-Gen (USB, Cleveland, OH), mutagênese direcionada ao local QuickChange

(Stratagene, San Diego, CA), mutagênese direcionada ao local mediada por PCR ou outros protocolos de mutagênese direcionados ao local.

Derivados

Os "derivados" incluem peptídeos, oligopeptídeos, polipeptídeos que podem, em comparação com a sequência de aminoácido da forma de ocorrência natural da proteína, tal como a proteína de interesse, compreende substituições de aminoácidos com resíduos de aminoácido que não ocorrem de maneira natural ou adições de resíduos de aminoácido que não ocorrem de maneira natural. Os "derivados" de uma proteína também abrangem peptídeos, oligopeptídeos, polipeptídeos que compreendem resíduos de aminoácido anterados de ocorrência natural (glicosilados, acilados, prenilados, fosforilados, miristoilados, sulfatados etc.) ou alterados que não ocorem de maneira natural em comparação com a sequência de aminoácido de uma forma de ocorrência natural do polipeptídeo. Um derivado também pode compreender um ou mais substituintes que não de aminoácido ou adições em comparação com a sequência de aminoácido a partir da qual este é derivado, por exemplo, uma molécula repórter ou oturo ligando, covalentemente ou não covalentemente ligado à sequência de aminoácido, tal como uma molécula repórter que é ligaa para facilitar sua detecção e resíduosd e aminoácido que não ocorrem de maneira natural com relação à sequência de aminoácido de uma proteína de ocorrência não natural. Além disso, os "derivados" também incluem fusões da forma de ocorrência natural da proteína com peptídeos de rotulação, tal como FLAG, HIS6 ou tioredoxina (para uma revisão de peptídeos de rotulação, ver Terpe, Appl. Microbiol. Biotechnol. 60, 523-533, 2003).

Ortólogo(s)/Parálogos(s)

Os ortólogos e parálogos abrangem conceitos evolucionários usados para descrever as relações ancestrais de genes. Os parálogos são genes dentro da mesma espécie que se originaram através da duplicação de um gene

ancestral; os ortólgos são genes de organismos diferentes que se originaram através da especialização e também são derivados de um gene ancestral comum.

Domínio

O termo "domínio" refere-se a uma série de aminoácidos conservados em posições específicas junto com o alinhamento de sequências de proteínas evolucionariamente relacionadas. Enquanto os aminoácidos em outras posições podem variar entre os homólogos, os aminoácidos que são altamente conservados em posições específicas indicam aminoácidos que são provavelmente essenciais na estrutura, estabilidade ou função de uma proteína. Identificado por seu alto grau de conservação em sequências alinhadas de uma família de homólogos de proteína, estes podem ser usados como identificadores para determinar se qualquer polipeptídeo em questão pertene a uma família de polipeptídeo previamente identificada.

Motivo/Sequência de Consenso/Assinatura

O termo "motivo" ou "sequência de consenso" ou "assinatura" refere-se a uma região conservada curta na sequência de proteínas evolutionariamente relacionadas. Os motivos são frequentemente partes altamente conservadas de domínios, mas também podem incluir apenas parte do domínio ou estarem localizados fora do domínio conservado (se todos os aminoácidos do motivo estiverem fora de um domínio definido).

Hibridização

O termo "hibridização" como definido neste é um processo em que as sequências de nucleotídeo complementares substancialmente homólogas fortalecem umas às oturas. O processo de hibridização pode ocorrer totalmente na solução, isto é, ambos os ácidos nucleicos complementares estão na solução. O processo de hibridização também pode ocorrer com um dos ácidos nucleicos complementares imobilizados a uma matriz, tal como gotas magnéticas, gotas de Sepharose ouqualquer outra

resina. Além disso, o processo de hibridização pode ocorrer com um dos ácidos nucleicos complementares imobilizados em um suporte sólido, tal como uma nitro-celulose ou membrana de náilon ou imobilizado, por exemplo, por fotolitografia, por exemplo, a um suporte de vidro siliceoso (o último conhecido como séries ou microsséries de ácido nucleico ou como aparas de ácido nucleico). A fim de permitir que a hibridização ocorra, as moléculas de ácido nucleico são, no geral, térmica ou quimicamente denaturadas para a fusão de um filamento duplo em dois filamentos simples e/ou para remover grampos de cabelo ou outras estruturas secundárias de ácidos nucleicos de filamento simples.

O termo "estringência" refere-se às condições sob as quais uma hibridização acontece. A estringência de hibridização é influenciada pelas condições, tais como temperatura, concentração de sal, força iônica e composição de tampão de hibridização. No geral, as condições de estringência baixa são selecionadas serem de cerca de 30° C menor do que o ponto de fusão térmico (Tm) para a sequência específica em uma força iônica e pH definidos. As condições de estringência médias ocorrem quando a temperatura está 20° C abaixo da Tm e as condições de estringência alta ocorrem quando a temperatura alta está 10° C abaixo da Tm. As condições de hibridização de estringência alta são tipicamente suadas para o isolamento de sequências de hibridização que têm similaridade de sequência alta com a sequência de ácido nucleico alvo. Entretanto, os ácidos nucleicos podem desviar na sequência e ainda codificar um polipeptídeo substancialmente idêntico, devido à degeneração do código genético. Portanto, as condições de hibridização de estringência média podem ser, algumas vezes, necessárias para identificar tais moléculas de ácido nucleico.

A Tm é a temperatura sob a força iônica definida e o pH em que 50 % da sequência alvo hibridiza-se a uma sonda perfeitamente equiparada. A Tm é dependente das condições de solução e a composição de

base e o comprimento da sonda. Por exemplo, as sequências mais longas hibridizam-se especificamente em temperaturas mais altas. A taxa máxima de hibridização é obtida em torno de 16° C até 32° C abaixo da Tm. A presença de cátions monovalentes na solução de hibridização reduz a repulsão eletrostática entre os dois filamentos de ácido nucleico, desse modo, promovendo a formação de híbrido; este efeito é visível para as concentrações de sódio de até 0,4 M (para concentrações mais altas, este efeito pode ser ignorado). A formamida reduz a temperatura de fusão de duplexes de DNA-DNA e de DNA-RNA com 0,6 a 0,7° C para cada percentual de formamida e adição de 50 % de formamida permite que a hibridização seja realizada de 30 a 45° C, embora a taxa de hibridização seja diminuída. As má combinações de pares de base reduzem a taxa de hibridização e a estabilidade térmica dos duplexes. Em média e para as sondas grandes, a Tm diminui cerca de 1° C por % de má combinação de base. A Tm pode ser calculada usando-se as seguitnes equações, dependendo dos tipos de híbridos:

1) híbridos de DNA-DNA (Meinkoth and Wahl, Anal. Biochem., 138: 267-284, 1984):

$Tm = 81,5° C + 16,6x\log_{10}[Na^+]^a + 0,41x \%[G/C^b] - 500x[L^C]^{-1} - 0,61x \%$ de formamida

2) híbridos de DNA-RNA ou RNA-RNA:

$Tm = 79,8 + 18,5 (\log_{10}[Na^+]^a) + 0,58 (\%G/C^b) + 11,8 (\%G/C^b)^2 - 820/L^C$

3) híbridos de oligo-DNA ou oligo-RNA[d]:

para <20 nucleotídeos: $Tm = 2 (I_n)$

para 20 a 35 nucleotídeos: $Tm = 22 + 1,46 (I_n)$

[a] ou para outro cátion monovalente, mas apenas preciso na faixa de 0,01 a 0,4 M.

[b] apenas preciso para % de GC na faixa de 30 % a 75 %.

[c] L = Comprimento de duplex em pares de base.

[d] oligo, oligonucleotídeo; I_n, = comprimento eficaz de iniciador = 2 x (n° de G/C) + (n° de A/T).

A ligação não específica pode ser controlada usando-se qualquer uma de diversas técnicas conhecidas tais como, por exemplo, O bloqueio de membrana com soluções contendo proteína, adições de RNA, DNA e SDS heterólogos ao tampão de hibridização e tratamento com Rnase. Para sondas não homólogas, uma série de hibridizações pode ser reealizada pela variação de um de (i) diminuir progressivamente a temperatura de recozimento (por exemplo, de 68° C a 42° C) ou (ii) diminuir progressivamente a concentração de formamida (por exemplo, de 50 % a 0 %). O técnico habilitado estará ciente de vários parâmetros que podem ser alterados durante a hibridização e que manterão ou mudarão as condições de estringência.

Além das condições de hibridização, a especificidade de hibridização tipicamente também dependerá da função de lavagens pós-hibridização. Para remover fudnamento que resultam de hibridização não específica, as amostras são lavadas com soluções salinas diluídas. Os fatores críticos de tais lavagens incluem a força iônica e a temperatura da solução de lavagem final: quanto menor a concentração de sal mais alta a temperatura de lavagem, mais alta a estringência da lavagem. As condições de lavagem são tipicamente realizadas em ou abaixo da estringência de hibridização. Uma hibridização positiva dá dá um sinal que é pelo menos duas vezes aquele dos fundamentos. No geral, as condições de estringência adequadas para os ensaios de hibridização de ácido nucleico ou procedimentos de detecção de amplificação do gene são apresentadas acima. Condições mais ou menos estringentes também podem ser selecionadas. O técnico habilitado está ciente de vários parâmetros que podem ser alterados durante a lavagem e que manterá ou mudará as condições de estringência.

Por exemplo, as condições de hibridização de alta estringência

típicas para os híbridos de DNA mais longas do que 50 nucleotídeos abrange a hibridização a 65° C em 1 x SSC ou em 42° C em 1 x SSC e 50 % de formamida, seguido pela lavagem a 65° C em 0,3 x SSC. Os exemplos de condições de hibridização de estringência média para híbridos de DNA mais longos do que 50 nucleotídeos abrange a hibridização a 50° C em 4 x SSC ou em 40° C em 6 x SSC e 50 % de formamida, seguido pela lavagem a 50° C em 2 x SSC. O comprimentodo híbrido é o comprimento antecipado para a hibridização de ácido nucleico. Quando os ácidos nucleicos de sequência conhecida são hibridizados, o comprimento do híbrido pode ser determinado pelo alinhamento das sequências e identificação das regiões conservadas descritas neste. 1 x SSC é 0,15 M de NaCl e 15 mM de citrato de sódio; a solução de hibridização e as soluções de lavagem podem, adicionalmente, incluir 5x reagente de Denhardt, 0,5 a 1,0 % de SDS, 100 μg/ml de DNA de esperma de salmão fragmentado desnaturado, 0,5 % de pirofosfato de sódio.

Para os propósitos de definir o nível de estringência, referência pode ser feita a Sambrook et al. (2001) Molecular Cloning: a laboratory manual, Edition, Cold Spring Harbor Laboratory Press, CSH, New York ou a Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989 e atualizações anuais).

Variante de união

O termo "variante de união" como usado neste abrange variantes de uma sequência de ácido nucleico em que os íntrons e/ou os exons selecionados foram excisados, trocados, substituídos ou adicionados ou em que os íntrons foram encurtados ou alongados lengthened. Tais variantes serão umas em que a atividade biológica da proteína é substancialmente retida; isto pode ser atingido retendo-se seletivamente os segmentos funcionais da proteína. Tais variantes de união podem ser encontradas em estado natural ou pode ser feito pelo homem. Os métodos para predizer e isolar tais variantes de união são bem conhecidos na técnica (ver, por

exemplo, Foissac and Schiex (2005) BMC Bioinformatics 6: 25).

Variante alélica

Os alelos e as variantes alélicas são formas alternativas de um dado gene, localizado na mesma posição cromossômica. As variantes alélicas abrangem Polimorfismos de Nucleotídeo Simples (SNPs), bem como Polimorfismos de Inserção/Anulação pequenos (INDELs). O tamanho de INDELs é usualmente menor do que 100 bp. SNPs e INDELs formam op grupo maior de variantes de sequência em cepas polimorficas de ocorrência natural of most organisms.

Embaralhamento de gene/Evolução direcionada

O embaralhamento de gene ou a evolução direcionada consistem de interações de embaralhamento de DNA seguido pela avaliação e/ou seleção apropriadas para gerar variantes de ácidos nucleicos ou porções destes que codificam proteínas tendo uma atividade biológica modificada (Castle et al., (2004) Science 304(5674): 1151-4; Patentes US 5.811.238 e 6.395.547).

Elemento regulador/Sequência de controle/Promotor

Os termos "elemento regulador", "sequência de controle" e "promotor" são todos usados de maneira intercambeável aqui e devem ser entendidos no contexto amplo para referir-se às sequências de ácido nucleico reguladoras capazes de realizar a expressão das sequências às quais estes estão ligados. O termo "promotor" tipicamente refere-se a uma sequência de controle de ácido nucleico localizada a montante do início de transcrição de um gene e que está envolvido no reconhecimento e ligação da RNA polimerase e outras proteínas, desse modo direcionando a transcrição de um ácido nucleico operacionalmente ligado. Abrangidos pelos termos já mencionados estão as sequências reguladoras de transcrição derivadas de um gene genômico eucariótico clássico (incluindo a caixa TATA que é requerida para a iniciação de transcrição precisa, with or without a CCAAT box

sequence) e elementos reguladores adicionais (isto é, sequências ativadoras a montante, intensificadores e silenciadores) que altera a expressão genética em resposta aos estímulos de desenvolvimento e/ou externos ou de um maneira específica de tecido. Também incluído dentro do termo está uma sequência reguladora de transcrição de um gene procariótico clássico, em cada caso este pode incluir uma sequência de -35 compartimentos e/ou sequências reguladoras de transcrição de -10 compartimentos. O termo "elemento regulador" também abrange uma molécula de fusão sintética ou derivado que confere, ativa ou intensifica a expressão de uma molécula de ácido nucleico em uma célula, tecido ou órgão.

A "promotor vegetal" compreende elementos reguladores, que mediam a expressão de um segmento de sequência codificadora em células de planta. Consequentemente, um promotor vegetal não necessita ser de origem vegetal, mas pode originar-se de vírus ou micro-organismos, por exemplo, de vírus que atacam células de planta. O "promotor vegetal" também pode originar-se de uma célula de planta, por exemplo, da planta que é transformada com a sequência de ácido nucleico a ser expressada no processo inventivo e descrita neste. Isto também aplica-se a outros sinais reguladores de "planta", tais como terminadores de "planta". Os promotores a montante das sequências de nucleotídeo úteis nos métodos da presente invenção podem ser modificados por uma ou mais substituições, insersões e/ou anulações de nucleotídeo sem interferir com a functionalidade ou atividade dos promotores, a estrutura de leitura aberta (ORF) ou a região reguladora 3', tal como terminadores ou outras regiões reguladoras 3' que estão localizadas além do ORF. Além disso, é possível que a atividade dos promotores seja aumentada por modificação de sua sequência ou que seja completamente substituído por promotores mais ativos, mesmo promotores de organismos heterólogos. Para a expressão em plantas, a molécula de ácido nucleico, como descrito acima, deve ser ligado operacionalmente ou compreender um promotor adequado

que expressa o gene no ponto certo no tempo e com o padrão de expressão espacial requerido.

Para a identificação de promotores funcionalmente equivalentes, a força do promotor e/ou o padrão de expressão de um promotor candidato pode ser analisada, por exemplo, ligando-se operacionalmente o promotor a um gene repórter e analisando-se o nível de expressão e o padrão do gene repórter em vários tecidos da planta. Os genes repórteres bem conhecidos adequados inclue, por exemplo, beta-glucuronidase ou beta-galactosidase. A atividade de promotor analisada medindo-se a atividade enzimática da beta-glucuronidase ou beta-galactosidase. A força do promotor e/ou padrão de expressão podem ser então comparadas àquela de um promotor de referência (tal como o usado nos métodos da presente invenção). Alternativamente, a força do promotor pode ser estimada quantifincando-se os níveis de mRNA ou comparando-se os níveis de mRNA do ácido usado nos métodos da presente invenção, com os níveis de mRNA de genes de manutenção, tais como 18S rRNA, usando-se os métodos conhecidos na técnica, tais como Northern blotting com análise densitométrica de autorradiogramas, PCR em tempo real quantitativo PCR ou RT-PCR (Heid et at., 1996 Genome Methods 6: 986-994). No geral, por "promotor fraco" é pretendido que um promotor conduza a expressão de uma sequência codificadora em um nível baixo. Por "nível baixo" são pretendidos níveis de cerca de 1/10.000 transcrições a cerca de 1/100.000 transcrições a cerca de 1/500.0000 transcrições por célula. Inversamente, um "promotor forte" conduz a expressão de uma sequência codificadora em nível alto ou em torno de 1/10 transcrições a cerca de 1/100 transcrições a cerca de 1/1000 transcrições por célula. No geral, por "promotor de força média" é pretendido um promotor que conduza a expressão de uma sequência codificadora em um nível menor do que um promotor forte, em particular, em um nível que está em todos os exemplos abaixo daqueles obtidos quando sob o controle de um

promotor 35S CaMV.

Operacionalmente ligado

O termo "operacionalmente ligado" como usado neste refere-se à ligação funcional entre a sequência de promotor e o gene de interesse, tal que a sequência de promotor seja capaz de iniciar a transcrição do gene de interesse.

Promotor constitutivo

Um "promotor constitutivo" refere-se a um promotor que é trasncripcionalmente ativo durante a maioria, mas não necessariamente, todas as fases de crescimento e desenvolvimento e sob a maioria das condições ambientais, em pelo menos uma célula, tecido ou órgão. A Tabela 2a abaixo dá exemplos de promotores constitutivos.

Tabela 2a: Exemplos de promotores constitutivos

Fonte de Gene	**Referência**
Actina	McElroy et al, Plant Cell, 2: 163-171, 1990
HMGP	WO 2004/070039
CAMV 35S	Odell et al, Nature, 313: 810-812, 1985
CaMV 19S	Nilsson et al., Physiol. Plant. 100:456-462, 1997
GOS2	de Pater et al, Plant J Nov;2(6):837-**44,** 1992, WO 2004/065596
Ubiquitina	Christensen et al, Plant Mol. Biol. 18: 675-689, 1992
Ciclofilina de arroz	Buchholz et al, Plant Mol Biol. 25(5): 837-**43,** 1994
Hidtona de milho H3	Lepetit et al, Mol. Gen. Genet. 231:276-285, 1992
Histona de alfafa H3	Wu et al. Plant Mol. Biol. 11:641-649, 1988
Actina 2	An et al, Plant J. 10(1); 107-121, 1996
34S FMV	Sanger et al., Plant. Mol. Biol., 14, 1990: 433-443
Rubisco small subunit	US 4,962,028
OCS	Leisner (1988) Proc Natl Acad Sci USA 85(5): 2553
SAD1	Jain et al., Crop Science, 39 (6), 1999: 1696
SAD2	Jain et al., Crop Science, 39 (6), 1999: 1696
nos	Shaw et al. (1984) Nucleic Acids Res. 12(20):7831-7846
V-ATPase	WO 01/14572
Super promotor	WO 95/14098
Proteínas G-box	WO 94/12015

Promotor ubíquo

Um promotor ubíquo é ativo substancialmente em todos os tecidos ou celuals de um organismo.

Promotor regulado por desenvolvimento

Um promotor regulado por desenvolvimento é ativo durante certos estágios de desenvolvimento ou em partes da planta que sofrem mudanças de desenvolvimento.

Promotor indutível

Um promotor indutível induziu ou aumentou a iniciação de transcrição em resposta a um estímulo químico, ambiental ou físico (para uma revisão ver Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48:89-108), ou pode ser "indutível por tensão", isto é, ativado quando a planta é exposta a várias condições de tensão ou uma "indutível por patogênio" isto é, ativado quando uma planta é exposta a vários patogênios.

Promotor específico de órgão/específico de tecido

Um promotor específico de órgão ou específico de tecido é um que é capaz de iniciar preferivelmente a transcrição em certos órgãos ou tecidos, tais como as folhas, raízes, tecido de semente etc. por exemplo, um "promotor específico de raiz" é um promotor que é transcripcionalmente ativo predominantemente em raízes vegetais, substancialmente para a exclusão de quaisquer outras partes de uma planta, enquanto ainda permite-se qualquer expressão fraca nestas outras partes de plantas. Os promotores capazes de iniciar a transcrição em certas células apenas são referidas neste como "específico de célula".

Os exemplos de promotores específicos de raiz são listadas na Tabela 2b abaixo:

Tabela 2b: Exemplos de promotores específicos de raiz

Fonte de Gene	**Referência**
RCc3	Plant Mol Biol. 1995 Jan;27(2):237-48
Arabidopsis PHT1	Kovama et al., 2005; Mudge et al. (2002, Plant *J.* 31:341)
Transportador de fosfato de Medicago	Xiao et al., 2006

Arabidopsis Pyk10	Nitz et al. (2001) Plant Sci 161(2): 337-346
genes expressiveis em raiz	Tingey et al., EMBO J. 6: 1, 1987.
genes indutíveis por auxina de tabaco	Van der Zaal et al., Plant Mol. Biol. 16, 983, 1991.
-tubulina	Oppenheimer, et al., Gene 63: 87, 1988.
Genes especificos da raiz de tabaco	Conkling, et al., Plant Physiol. 93: 1203, 1990.
Gene G1-3b de B. napus	United States Patent No. 5, 401, 836
SbPRP1	Suzuki et al., Plant Mol. Biol. 21: 109-119, 1993.
LRX1	Baumberger et al. 2001, Genes & Dev. 15:1128
BTG-26 Brassica napus	US 20050044585
LeAMT1 (tomate)	Lauter et al. (1996, PNAS 3:8139)
The LeNRT1-1 (tomate)	Lauter et al. (1996, PNAS 3:8139)
gene de patatina classe I (batata)	Liu et al., Plant Mol. Biol. 153:386-395, 1991.
KDC1 (Daucus carota)	Downey et al. (2000, J. Biol. Chem. 275:39420)
Gene TobRB7	W Song (1997) PhD Thesis, North Carolina State University, Raleigh, NC USA
OsRAB5a (Arroz)	Wang et al. 2002, Plant Sci. 163:273
ALF5 (Arabidopsis)	Diener et al. (2001, Plant Cell 13:1625)
NRT2;1 Np (N. plumbaginifolia)	Quesada et al. (1997, Plant Mol. Biol. 34:265)

Um promotor específico de semente é transcricionalmente ativo predominantemente no tecido da semente, mas não exclusivamente necessário no tecido da semente (em casos de expressão fraca). O promotor específico de semente pode ser ativo durante o desenvolvimento e/ou durante a germinação de semente. O promotor específico de semente pode ser específico de endosperma/aleurona/embrião. Os exemplos de promotor específico de sementes (específico de endosperma/aleurona/embrião) sãom ostrados na Tabela 2d, 2e, 2f. Ainda, os exemplos de promotores específicos de semente são dados em Qing Qu and Takaiwa (Plant Biotechnol. J. 2, 113-125, 2004), cuja divulgação é incorporada por referência neste como se totalmente apresentado.

Tabela 2c: Exemplos de promotor específico de sementes

gene semelhante à catepsina β - cevada Ltp2 – milho B-Peru Selinger et at., Genetics 149;1125-38,1998

Fonte de Gene	**Referência**
genes específicos de semente	Simon et al., Plant Mol. Biol. 5: 191, 1985;
	Scofield et al., J. Biol. Chem. 262: 12202, 1987.;
	Baszczynski et al., Plant Mol. Biol. 14: 633, 1990.
albumina de castanha-do-	Pearson et al., Plant Mol. Biol. 18: 235-245, 1992.

pará	
legumina	Ellis et al., Plant Mol. Biol. 10: 203-214, 1988.
glutelina (arroz)	Takaiwa et al., Mol. Gen. Genet. 208: 15-22, 1986;
	Takaiwa et al., FEBS Letts. 221: 43-47, 1987.
zeína	Matzke et al Plant Mol Biol, 14(3):323-32 1990
napA	Stalberg et al, Planta 199: 515-519, 1996.
trigo LMW e HMW glutenin-1	Mol Gen Genet 216:81-90, 1989; NAR 17:461-2, 1989
trigo SPA	Albani et al, Plant Cell, 9: 171-184, 1997
α, β, γ-gliadinas de trigo	EMBO J. 3:1409-15, 1984
promotor Itr1 de cevada	Diaz et aI. (1995) Mol Gen Genet 248(5):592-8
cevada B1, C, D, hordeína	Theor Appl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750-60, 1996
cevada DOF	Mena et al, The Plant Journal, 116(1): 53-62, 1998
blz2	EP99106056.7
promotor sintético	Vicente-Carbajosa et al., Plant J. 13: 629-640, 1998.
prolamina de arroz NRP33	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
a-globulina de arroz Glb-1	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
arroz OSH1	Sato et al, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996
a-globulina de arroz REB/OHP-1	Nakase et al. Plant Mol. Biol. 33: 513-522, 1997
ADP-glicose pirofosforilase de arroz	Trans Res 6:157-68, 1997
família do gene ESR do milho	Plant J 12:235-46, 1997
α-cafirina de sorgo	DeRose et al., Plant Mol. Biol 32:1029-35, 1996
KNOX	Postma-Haarsma et al, Plant Mol. Biol. 39:257-71, 1999
oleosina de arroz	Wu et al, J. Biochem. 123:386, 1998
oleosina de girassol	Cummins et al., Plant Mol. Biol. 19: 873-876, 1992
proteína ribossômica 40S de arroz putativo	WO 2004/070039
alanina aminotransferase de arroz	não publicado
inibidor de tripsina ITR1 (cevada)	não publicado
PRO0151, WSI18 de arroz	WO 2004/070039
PRO0175, RAB21 de Arroz	WO 2004/070039
PR0005	WO 2004/070039
PR00095	WO 2004/070039
α-amilase (Amy32b)	Lanahan et al, Plant Cell 4:203-211, 1992; Skriver et al, Proc Natl Acad Sci USA 88:7266-7270, 1991
cathepsin R-like gene	Cejudo et al, Plant Mol Biol 20:849-856, 1992
Barley Ltp2	Kalla et al., Plant J. 6:849-60, 1994
Chi26	Leah et al., Plant J. 4:579-89, 1994
Maize B-Peru	Selinger et al., Genetics 149;1125-38,1998

Tabela 2d: exemplos de promotores específicos de endosperma

Fonte de Gene	Referência
glutelina (arroz)	Takaiwa et al. (1986) Mol Gen Genet 208:15-22; Takaiwa et al. (1987) FEBS Letts. 221:43-47
zeína	Matzke et al., (1990) Plant Mol Biol 14(3): 323-32
trigo LMW e HMW glutenin-1	Colot et al. (1989) Mol Gen Genet 216:81-90, Anderson et al. (1989) NAR 17:461-2
trigo SPA	Albani et al. (1997) Plant Cell 9:171-184
gliadinas de trigo	Rafalski et al. (1984) EMBO 3:1409-15
promotor Itr1 de cevada	Diaz et al. (1995) Mol Gen Genet 248(5):592-8
cevada B1, C, D, hordeína	Cho et al. (1999) Theor Appl Genet 98:1253-62; Muller et al. (1993) Plant J 4:343-55; Sorenson et al. (1996) Mol Gen Genet 250:750-60
cevada DOF	Mena et al, (1998) Plant J 116(1): 53-62
blz2	Onate et al. (1999) J Biol Chem 274(14):9175-82
promotor sintético	Vicente-Carbajosa et al. (1998) Plant J 13:629-640
prolamina de arroz NRP33	Wu et al, (1998) Plant Cell Physiol 39(8) 885-889
rice globulin Glb-1	Wu et al. (1998) Plant Cell Physiol 39(8) 885-889
globulina de arroz REB/OHP-1	Nakase et al. (1997) Plant Molec Biol 33: 513-522
ADP-glicose pirofosforilase de arroz	Russell et al. (1997) Trans Res 6:157-68
família do gene ESR do milho	Opsahl-Ferstad et al. (1997) Plant J 12:235-46
cafirina de sorgo	DeRose et al. (1996) Plant Mol Biol 32:1029-35

Tabela 2e: Exemplos de promotores específicos de embrião:

Fonte de Gene	Referência
arroz OSH1	Sato et al, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996
KNOX	Postma-Haarsma et al, Plant Mol. Biol. 39:257-71, 1999
PRO0151	WO 2004/070039
PRO0175	WO 2004/070039
PRO005	WO 2004/070039
PRO0095	WO 2004/070039

Tabela 2f: Exemplos de promotores específicos de aleurona:

Fonte de Gene	Referência
α-amilase (Amy32b)	Lanahan et al, Plant Cell 4:203-211, 1992; Skriver et al, Proc Natl Acad Sci USA 88:7266-7270, 1991
gene semelhante à catepsina β	Cejudo et al, Plant Mol Biol 20:849-856, 1992
cevada Ltp2	Kalla et al., Plant J. 6:849-60, 1994
Chi26	Leah et al., Plant J. 4:579-89, 1994
milho B-Peru	Selinger et al., Genetics 149;1125-38,1998

Um promotor específico de tecido verde como definido aqui é um promotor que é transcricionalmente ativo predominantemente em tecido verde, substancialmente para a inclusão de quaisquer outras partes de uma planta, enquanto ainda permite-se qualquer expressão fraca nestas outras

partes de planta.

Os exemplos de promotores específicos de tecido verde que podem ser usados para realizar ps métodos da invenção são mostrados na Tabela 2g.

Tabela 2g: Exemplos de promotores específicos de tecido verde

Gene	Expressão	Referência
Ortofosfato diquinase de milho	específico de folha	Fukavama et al., 2001
Fosfoenolpiruvato carboxilase de milho	específico de folha	Kausch et al., 2001
Fosfoenolpiruvato carboxilase de arroz	específico de folha	Liu et al., 2003
subunidade pequena de arroz Rubisco	específico de folha	Nomura et al., 2000
beta expansina de arroz EXBP9	específico de broto	WO 2004/070039
subunidade pequena de Pigeonpea Rubisco	específico de folha	Panguluri et al., 2005
Ervilha RBCS3A	específico de folha	

Um outro exemplo de um promotor específico de tecido é um promotor específico de meristema, que é transcricionalmente ativo predominantemente em tecido meristemático, substancialmente para a xclusão de quaisquer outras partes de uma planta, enquanto ainda permite-se qualquer expressão fraca nestas outras partes vegetais. Os exemplos de específicos de meristema verde que podem ser usados para realizar os métodos da invenção são mostrados na Tabela 2h abaixo.

Tabela 2h: Exemplos de promotores específicos de meristema

Fonte de Gene	Padrão de expressão	Referência
arroz OSH1	Meristema apical de broto, de estágio globular de broto ao estágio de muda	Sato *et al.* (1996) Proc. Natl. Acad. Sci. USA, 93: 8117-8122
Metalotioneína de arroz	Específico de meristema	BAD87835.1
WAK1 & WAK 2	Meristemas apicais de broto e raiz apical e na expansão de folhas e sépalas	Wagner & Kohorn (2001) Plant Cell 13(2): 303-318

Terminador

O termo "terminador" abrange uma sequência de controle que é uma sequência de DNA no final de uma unidade transcripcional cujo processamento dos sinais 3' e a poliadenilação de uma transcrição primária e terminação de transcrição. O terminador pode ser derivado do gene natural, de uma variedade de outros genes vegetais ou de T-DNA. O terminador a ser

adicionado pode ser derivado, por exemplo, dos genes de nopalina sintase ou octopina sintase ou alternativamente de um outro gene vegetal ou menos preferivelmente de qualquer outro gene eucariótico.

Modulação

O termo "modulação" significa em relação à expressão ou expressão genética, um processo em que o nível de expressão é mudado pela dita expressão de gene em comparação com a planta de controle, o nível de expressão pode ser aumentado ou diminuído. A expressão não modulada original pode ser de qualquer tipo de expressão de um RNA (rRNA, tRNA) ou mRNA estruturais com tradução subsequente. O termo "modular a atividade" deve significar qualquer mudança da expressão das sequências de ácido nucleico ou proteínas codificadas inventivas, que leva ao rendimento aumentado e/ou desenvolvimento aumentado das plantas.

Expressão

O termo "expressão" ou "expressão de gene" significa a transcrição de um gene específico ou genes específicos ou construção genética específica. O termo "expressão" ou "expressão de gene" em particular, significa a transcrição de um gene ou genes ou construção genética em RNA (rRNA, tRNA) ou mRNA estruturais com ou sem a tradução subsequente do último em uma proteína. O processo inclui a transcrição de DNA e processamento do produto de mRNA resultante.

Expressão aumentada/superexpressão

O termo "expressão aumentada" ou "superexpressão" como usado neste significa qualquer forma de expressão que seja adicional ao nível de expressão do tipo selvagem original.

Os métodos para aumentar a expressão de genes ou produtos de gene são bem documentados na técnica e incluem, por exemplo, a superexpressão conduzida pelos promotores apropriados, o uso de intensificadores de transcrição ou intensificadores de tradução. Os ácidos

nucleicos isolados que servem como elementos promotores ou intensificadores podem ser introduzidos em uma posição apropriada (tipicamente a montante) de uma forma não heteróloga de um polinucleotídeo a fim de super-regular a expressão de um ácido nucleico que codifica o polipeptídeo de interesse. Por exemplo, os promotores endógenos podem ser alterados in vivo por mutação, anulação e/ou substituição (ver, Kmiec, US 5,565,350; Zarling et al., WO9322443) ou promotores isolados podem ser introduzidos em uma célula de planta na orientação apropriada em distância de um gene da presente invenção a fim de cotnrolar a expressão do gene.

Se a expressão polipeptídeo for desejada, é, no geral, desejável incluir uma região de poliadenilação na extremidade 3' de uma região codificadora de polinucleotídeo. A região de poliadenilação pode ser derivada do gene natural, de uma variedade de outros genes vegetais ou de T-DNA. A sequência de extremidade 3' a ser adicionada pode ser derivada, por exemplo, aos genes nopalina sintase ou octopina sintase ou, alternativamente de um outro gene vegetal ou menos preferivelmente de qualquer outro gene eucariótico.

Uma sequência de íntron também pode ser adicionada à região não traduzida 5' (UTR) ou à sequência codificadora da sequência codificadora parcial coding sequence para aumentar a quantidade da mensagem madura que acumula-se no citosol. A inclusão de um íntron unível na unidade de transcrição tanto nas construções de expressão vegetal ou animal foi mostrada aumentar a expressão genética tanto nos níveis de mRNA quanto de proteína de atpe 1000 vezes (Buchman and Berg (1988) Mol. Cell biol. 8: 4395-4405; Callis et al. (1987) Genes Dev 1:1183-1200). Tal intensificação de íntron de expressão genética é tipicamente maior do que quando colocada próximo à extremidade 5' da unidade de transcrição. O uso de íntrons de milho, íntron Adhl-S 1, 2, e 6, o íntron Bronze-1 são conhecidos na técnica. Para informação geral ver: The Maize Handbook, Capítulo 116, Freeling and

Walbot, Eds., Springer, N.Y. (1994).

Gene endógeno

Referência neste, um gene "endógeno" não refere-se apenas ao gene em questão como observado em uma planta em sua forma natural (isto é, sem haver qualquer intervenção humana), mas também refere-se àquele mesmo gene (ou um ácido nucleico/gene substancialmente homólogo) em uma forma isolada subsequentemente (re)introduzida em uma planta (um transgene). Por exemplo, a planta transgênica contendo um tal transgene pode encontrar uma redução substancial da expressão de transgene e/ou redução substancial de expressão do gene endógeno. O gene isolado pode ser isolado de um organismo ou pode ser feito pelo homem, por exemplo, pela síntese química.

Expressão diminuída

Referência aqui a "expressão diminuída" ou "redução ou eliminação substancial" de expressão é entendida significar uma diminuição na expressão do gene endógeno e/ou níveis de polipeptídeo e/ou atividade de polipeptídeo com relação às plantas de controle. A redução ou eliminação substancial está na ordem crescente de preferência pelo menos 10 %, 20 %, 30 %, 40 % ou 50 %, 60 %, 70 %, 80 %, 85 %, 90 % ou 95 %, 96 %, 97 %, 98 %, 99 % ou mais reduzido em comparação com aqueles de plantas de controle.

Para a redução ou eliminação substancial de expressão de um gene endógeno em uma planta, um comprimento suficiente de nucleotídeos substancialmente contíguos de uma sequência de ácido nucleico é requerido. A fim de realizar o silencimento de gene, isto pode ser tão pouco quanto 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10 ou menos nucleotídeos, alternativamente isto pode ser tanto quanto o gene total (incluindo o 5' e/ou 3' UTR, em parte em todo). A extensão de nucleotídeo substancialmente contíguo pode ser derivada do ácido nucleico que codifica a proteína de interesse (gene alvo) ou

de qualquer ácido nucleico capaz de codificar um ortólogo, parálogo ou homólogo da proteína de interesse. Preferivelmente, a tensão de nucleotídeos substancialmente contíguos é capaz de formar ligações de hidrogênio com o gene alvo (filamento sentido ou anti-sentido), mais preferivelmente, a tensão de nucleotídeos substancialmente contíguos tem, na ordem crescente de preferência, 50 %, 60 %, 70 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 %, 100 % de identidade de sequência ao gene alvo (filamento sentido ou anti-sentido). Uma sequência de ácido nucleico que codifica um polipeptídeo (funcional) não é um requerimento para os vários métodos debatidos neste para a redução ou eliminação substancial de expressão de um gene endógeno.

Esta redução ou eliminação substancial de expressão pode ser atingida usando-se ferramentas e técnicas de rotina. Um método preferido para a redução ou eliminação substancial de expressão de gene endógeno ocorre pela introdução e expressão em uma planta para a construção genética em que o ácido nucleico (neste caso, uma extensão de nucleotídeos substancialmente contíguos derivados do gene de interesse ou de qualquer ácido nucleico capaz de codificar um ortólogo, parálogo ou homólogo de qualquer da proteína de interesse) é clonado como uma repetição invertida (em parte ou completamente), separado por um espaçador (DNA não codificador).

Em um tal método preferido, a expressão do gene endógeno é reduzida ou substancialmente eliminada através do silenciamento mediado por RNA usando-se uma repetição invertida de um ácido nucleico ou uma parte destes (neste caso uma extensão de nucleotídeos substancialmente contíguos derivados do gene de interesse ou de qualquer ácido nucleico capaz de codificar um ortólogo, parálogo ou homólogo da proteína de interesse), preferivelmente capaz de formar uma estrutura de grampo de cabelo. A repetição invertida é clonada em um vetor de expressão que compreende sequências de controle. Uma sequência de DNA não codificadora de ácido

nucleico (um espaçador, por exemplo, um fragmento de região de ligação de matriz (MAR), um íntron, um poliligador etc.) está localizado entre os dois ácidos nucleicos invertidos que formam a repetição invertida. After transcription of the inverted repeat, a chimeric RNA com a self-complementary structure is formed (partial or complete). Esta estrutura de RNA de filamento duplo é referida como o RNA grampo de cabelo (hpRNA). O hpRNA é processado pela planta em siRNAs que são incorporados em um complexo de silenciamento induzido por RNA (RISC). O RISC ainda cliva as transcrições de mRNA, desse modo reduzindo substancialmente o número de transcrições de mRNA a serem traduzidas em polipeptídeos. Para detalhes gerais adicionais ver, por exemplo,, Grierson et al. (1998) WO 98/53083; Waterhouse et al. (1999) WO 99/53050).

O desempenho dos métodos da invenção não contam com a introdução e com a expressão em uma planta para a construção genética em que o ácido nucleico é clonado como uma repetição invertida, mas qualquer um ou mais de diversos métodos "silenciadores de gene" bem conhecidos podem ser usados para atingir os mesmos efeitos.

Um tal método para a redução da expressão de gene endógeno é o silenciamento mediado por RNA de expressão genética (sub-regulação). O silenciamento neste caso é disparado em uma planta por uma sequência de RNA de filamento duplo (dsRNA) que é substancialmente similar ao gene endógeno alvo. Este dsRNA ainda é processado pela planta de cerca de 20 a cerca de 26 nucleotídeos denominados RNAs de interferência curta (siRNAs). Os siRNAs são incorporados em um complexo silenciador induzido por RNA (RISC) que cliva a transcrição de mRNA do gene alvo endógeno, desse modo redunzindo substancialmente o número de transcrições de mRNA a serem traduzidas em um polipeptídeo. Preferivelmente, a sequência de RNA de filamento duplo corresponde a um gene alvo.

Um outro exemplo de um método de silenciamento de RNA

envolve a introdução de sequências de ácido nucleico ou partes destes (neste caso uma extensão de nucleotídeos substancialmente contíguos derivados do gene de interesse ou de qualquer ácido nucleico capaz de codificar um ortólogo, parálogo ou homólogo da proteína de interesse) em uma orientação sentido em uma planta. A "orientação sentido" refere-se à sequência de DNA que é homóloga a uma transcrição de mRNA desta. Introduzido em uma planta deve ser, portanto, pelos menos uma cópia da sequência de ácido nucleico. A sequência de ácido nucleico adicional reduzirá a expressão do gene endógeno, dando origem a um fenômeno conhecido como co-supressão. A redução de expressão de gene será mais pronunciada se diversas cópias adicionais de uma sequência de ácido nucleico forem introduzidas na planta, como existe uma correlação positiva entre níveis de transcrição altos e o disparo de co-supressão.

Um outro exemplo de método silenciador de RNA envolve o uso de sequências de ácido nucleico anti-sentido. Uma sequência de ácido nucleico "anti-sentido" compreende uma sequência de nucleotídeo que é complementar a uma sequência de ácido nucleico "sentido" que codifica um proteína, isto é complementar ao filamento codificador de uma molécula de cDNA de filamento duplo ou complementar a uma sequência de transcrição de mRNA. A sequência de ácido nucleico anti-sentido é preferivelmente complementar ao gene endógeno a ser silenciado. A complementaridade pode estar localizada na "região codificadora" e/ou na "região não codificadora" de um gene. O termo "região codificadora" refere-se a uma região da sequência do nucleotídeo que compreende códons que são traduzidos em resíduos de aminoácido. O termo "região não codificadora" refere-se às sequências 5' e 3' que flanqueiam a região codificadora que são transcritas mas não traduzidas em aminoácidos (também referidos como regiões não traduzidas 5' e 3').

As sequências anti-sentido de ácido nucleico podem ser projetadas de acordo com as regras da formação de pares de base de Watson

and Crick. A sequência de ácido nucleico anti-sentido pode ser complementar à sequência de ácido nucleico total (neste caso uma extensão de nucleotídeos substancialmente contíguos derivados do gene de interesse ou de qualquer ácido nucleico capaz de codificar um ortólogo, parálogo ou homólogo da proteína de interesse), mas também pode ser um oligonucleotídeo que é antisense a pelo menos uma parte da sequência de ácido nucleico (incluindo o mRNA 5' e 3' UTR). Por exemplo, a sequência de oligonucleotídeo anti-sentido pode ser complementar à região que circunda o local de início de tradução de uma transcrição de mRNA que codifica um polipeptídeo. O comprimento de uma sequência de oligonucleotídeoanti-sentido adequada é conhecido na técnica e pode começar a partir de cerca de 50, 45, 40, 35, 30, 25, 20, 15 ou 10 nucleotídeos de comprimento ou menos. Uma sequência de ácido nucleico anti-sentido de acordo com a invenção pode ser construída usando-se a síntese química e as reações de ligação enzimática usando-se métodos conhecidos na técnica. Por exemplo, uma sequência de ácido nucleico anti-sentido (por exemplo, uma sequência de oligonucleotídeo anti-sentido) pode ser quimicamente sintetizada usando-se nucleotídeos de ocorrência natural ou nucleotídeo variadamente modificados para aumentar a estabilidade biológica das moléculas ou para aumentar a estabilidade física do duplex formado entre as sequências de ácido nucleico anti-sentido e sentido, por exemplo, derivados de fosforotioato e nucleotídeos substituídos de acridina podem ser usados. Os exemplos de nucleotídeos que podem ser usados para as sequências sequências de ácido nucleico anti-sentido são bem conhecidos na técnica. As modificações de nucleotídeo conhecidas incluem metilação, ciclização e 'coberturas' e substituição de um ou mais dos nucleotídeos de ocorrência natural com uma análogo tal como inosina. Outras modificações de nucleotídeos são bem conhecidas na técnica.

A sequência de ácido nucleico anti-sentido pode ser produzida biologicamente usando-se um vetor de expressão em que uma sequência de

ácido nucleico foi subclonada em uma orientação anti-sentido (isto é, o RNA transcrito a partir do ácido nucleico inserido será de uma orientação anti-sentido a um ácido nucleico alvo de itneresse). Preferivelmente, a produção de sequências de ácido nucleico anti-sentido em plantas ocorre por meio de construção de ácido nucleico estavelmente integrada que compreende um promotor, um oligonucleotídeo anti-sentido operacionalmente ligado e um terminador.

As moléculas de ácido nucleico usadas para o silenciamento no método da invenção (se introduzidas em uma planta ou geradas in situ) hibridizam-se ou ligam-se com as transcrições de mRNA e/ou DNA genômico que codifica um polipeptídeo para, desse modo, inibir a expressão da proteína, por exemplo, pela inibição da transcrição e/ou tradução. A hibridização pode ocorrer pela complementaridade de nucleotídeo convencional para formar um duplex estável ou, por exemplo, no caso de uma molécula de ácido nucleico anti-sentido que liga-se aos duplexes de DNA, através de interações específicas no sulco principal da hélice dupla. As sequências anti-sentido de ácido nucleico podem ser introduzidas pela transformação ou pela injeção direta em um local de tecido específico. Alternativamente, as sequências de ácido nucleico anti-sentido podem ser modificadas paa alvejar células selecionadas e então administradas sistemicamente. Por exemplo, para a administração sistêmica, as sequências de ácido nucleico anti-sentido podem ser modificadas tal que estas liguem-se especificamente aos receptores ou antígenos expressados em uma superfície celular selecionada, por exemplo, pela ligação da sequência de ácido nucleico anti-sentido a peptídeos ou anticorpos que ligam-se aos receptores de superfície ou antígenos. As sequências anti-sentido de ácido nucleico também podem ser liberadas às células usando-se os vetores descritos neste.

De acordo com um outro aspecto, a sequência de ácido nucleico anti-sentido é uma sequência de ácido nucleico α-anomérica. Uma

sequência de ácido nucleico α-anomérica forma híbridos de filamento duplo específicos com RNA complementar em que, ao contrário das b-unidades usuais, os filamentos correm em paralelo um ao outro (Gaultier et al. (1987) Nucl Ac Res 15: 6625-6641). A sequência de ácido nucleico antisentido também pode compreender compreende um 2'-o-metilribonucleotídeo (Inoue et al. (1987) Nucl Ac Res 15, 6131-6148) ou um análogo de RNA-DNA quimérico (Inoue et al. (1987) FEBS Left. 215, 327-330).

A redução ou a eliminação substancial da expressão de gene endógeno também pode ser realizado usando-se ribozimas. As ribozimas são moléculas de RNA catalíticas com atividade de ribonuclease que são capazes de clivar uma sequência de ácido nucleico de filamento simples, tal como um mRNA, ao qual estes têm uma região complementar. Desta maneira, as ribozimas (por exemplo, ribozimas cabeça de martelo (descrito em Haselhoff and Gerlach (1988) Nature 334, 585-591) podem ser usadas para clivar cataliticamente as transcrições de mRNA que codificam um polipeptídeo, desse modo reduzindo substancialmente o número de transcrições de mRNA a serem traduzidas em um polipeptídeo. Uma ribozima tendo especificidade quanto a uma sequência de ácido nucleico pode ser designada (ver, por exemplo, Cech et al. Patente U. S. N° 4.987.071 e Cech et al. Patente U. S. N° 5.116.742). Alternativamente, as transcrições de mRNA que correspondem a uma sequência de ácido nucleico podem ser usadas para selecionar um RNA catalítico tendo uma atividade de ribonuclease específica de um grupo de moléculas de RNA (Bartel and Szostak (1993) Science 261, 1411-1418). O uso de ribozimas para o silenciamento de gene em plantas é conhecido na técnica (por exemplo, Atkins et al. (1994) WO 94/00012; Lenne et al. (1995) WO 95/03404; Lutziger et al. (2000) WO 00/00619; Prinsen et al. (1997) WO 97/13865 e Scott et al. (1997) WO 97/38116).

O silenciamento de gene também pode ser atingido pela mutagênese de inserção (por exemplo, inserção de T-DNA ou inserção de

transposon) ou por estratégias descritas por, entre outros, Angell and Baulcombe ((1999) Plant J 20(3): 357-62), (Amplicon VIGS WO 98/36083) ou Baulcombe (WO 99/15682).

O silenciamento do gene também pode ocorrer se houver uma mutação em um gene andógeno e/ou uma mutação em um gene/ácido nucleico isolados subsequentemente introduzido em uma planta. A redução ou a eliminação substancial podem ser causadas por um polipeptídeo não funcional. Por exemplo, o polipeptídeo pode ligar-se a várias proteínas de interação uma ou mais mutações e/ou truncações podem, portanto, fornecer um polipeptídeo que ainda é capaz de ligar-se às proteínas de interação (tais como proteínas receptoras) mas que não podem apresentar sua função normal (tal como ligando de sinalização).

Um método adicional para o silenciamento de gene é o alvejamento de sequências de ácido nucleico complementares à região reguladora do gene (por exemplo, o promotor e/ou intensificadores) para formar estruturas hélicas triplas que evitam a transcrição do gene em células alvo. Ver Helene, C., Anticancer Drug Res. 6, 569-84, 1991; Helene et al., Ann. N.Y. Acad. Sci. 660, 27-36 1992 e Maher, L.J. Bioassays 14, 807-15, 1992.

Outros métodos, tal como o uso de anticorpos direcionados a um polipeptídeo endógeno para inibir sua função na planta ou interferência no caminho de sinalização em que um polipeptídeo está envolvido, será bem conhecido pela pessoa habilitada. Em particular, pode ser considerado que as moléculas feitas pelo homem podem ser úteis para a inibição da função biológica de um polipeptídeo alvo ou para interferir com o caminho de sinalização em que o polipeptídeo alvo está envolvido.

Alternativamente, um programa de avaliação pode ser ajustado para identificar em uma população vegetal variantes naturais de um gene, cujas variantes codificam polipeptídeos com atividade reduzida. Tais

variantes naturais também podem ser usadas, por exemplo, para realziar a recombinação homóloga.

Os microRNAs artificiais e/ou naturais (miRNAs) podem ser usados para nocautear a expressão e/ou a tradução de mRNA. Os miRNAs endógenos são RNAs pequenos de filamento simples de tipicamente 19 a 24 nucleotídeos de comprimento. Estes funcionam primariamente para regular a expressão genética e/ou a tradução do mRNA. A maioria dos microRNAs vegetais (miRNAs) tem complementaridade perfeita ou quase perfeita com suas sequências alvo. Entretanto, existem alvos naturais com até cinco más combinações. Estes são processados a partir de RNAs não codificadores mais longo com estruturas de redobra características por RNases específicas de filamento duplo da família Dicer. No processamento, estes são incorporados no complexo de silenciamento induzido por RNA (RISC) pela ligação ao seu componente principal, uma proteína Argonaute. Os MiRNAs servem como os componentes de especificidade de RISC, visto que estes pares de base alvejam os ácidos nucleicos, a maioria dos mRNAs, no citoplasma. Os eventos reguladores subsequentes incluem clivagem de mRNA alvo e destruição e/ou inibição de tradução. Os efeitos da super-expressão de miRNA são, desta maneira, frequentemente refletidos em níveis de mRNA diminuídos de genes alvo.

Os microRNAs artificiais (amiRNAs), que são, tipicamente, 21 nucleotídeos de comprimento, podem ser geneticamente projetados especificamente para regular negativamente a expressão enética de genes simples ou múltiplos de interesse. Os determinantes de seleção alvo de microRNA vegetal são bem conhecidos na técnica. Os parâmetros empíricos para o reconhecimento alvo foram definidos e podem ser usados para auxiliar no projeto de amiRNAs específicos, (Schwab et al., Dev. Cell 8, 517-527, 2005). As ferramentas convenientes para o projeto e geração de amiRNAs e seus precursores também estão disponíveis ao público (Schwab et al., Plant

Cell 18, 1121-1133, 2006).

Para o desenpenho ótimo, as técnicas de silenciamento de gene usadas para reduzir a expressão em uma planta de um gene endógeno requer o uso de sequências de ácido nucleico de plantas monocotiledôneas para a transformação de plantas monocotiledôneas e de plantas dicotiledôneas para a transformação de plantas dicotiledôneas. Preferivelmente, uma sequência de ácido nucleico de uma dada espécie vegetal é introduzida naquela mesma espécie. Por exemplo, uma sequência de ácido nucleico de arroz é transformada em uma planta de arroz. Entretanto, não é um requerimento absoluto que a sequência de ácido nucleico a ser introduzida origine-se a partir das mesmas espécies vegetais como a planta em que este será introduzido. É suficiente que exista homologia substancial entre o gene alvo endógeno e o ácido nucleico a ser introduzido.

São descritos acima os exemplos de vários métodos para a redução ou eliminação de expressão em uma planta de um gene endógeno. Uma pessoa habilitada na técnica deve ser facilmente capaz de se adaptar os métodos já mencionados para o silenciamento a fim de atingir a redução da expressão de um gene endógeno em uma planta total ou em partes destes através do uso de um promotor apropriado, por exemplo.

Marcador Selecionável (gene)/Gene repórter

"Marcador selecionável", "gene de marcador selecionável" or "gene repórter" inclui qualquer gene que confira um fenotipo em uma célula em que esta é expressada para facilitar a identificação e/ou a seleção de células que são transfectadas ou transformadas com uma construção de ácido nucleico da invenção. Estes genes marcadores permitem a identificação de uma transferência bem sucedida das moléculas de ácido nucleico por intermédio de uma série de diferentes princípios. Os marcadores adequados podem ser selecionados de marcadores que conferem resistência antibiótica ou herbicida, que introduz um novo traço metabólico ou que permite a seleção

visual. Os exemplos de genes marcadores selecionáveis incluem genes que conferem resistência a antibióticos (tal como nptll que fosforila neomicina e canamicina ou hpt, que fosforila higromicina ou genes que conferem resistência a, por exemplo, bleomicina, estreptomicina, tetraciclina, cloranfenicol, ampicilina, gentamicina, geneticina (G418), espectinomicina ou blasticidina), a herbicidas (por exemplo, bar que fornece resistência a Basta®; aroA ou gox que fornece resistência contra glifosato ou os genes que conferem resistência a, por exemplo, imidazolinona, fosfinotricina ou sulfoniluréia) ou genes que fornecem um traço metabólico (tal como manA que permite que as plantas usem manose como a fonte de carbono única ou xilose isomerase para a utilização de xilose ou marcadores antinutritivos, tais como a resistência à 2-desoxiglicose). A expressão de genes marcadores resulta na formação de cor (por exemplo, β-glucuronidase, GUS ou β-galactosidase com seus substratos coloridos, por exemplo, X-Gal), luminescência (tal como o sistema de luciferina/luciferase) ou fluorescência (Proteína Fluorescente Verde, GFP e seus derivados). Esta lista representa apenas um número pequeno de marcadores possíveis. O trabalhador habilitado está familiarizado com tais marcadores. Os marcadores diferentes são preferidos, dependendo do organismo e do método de seleção.

É conhecido que na integração estável ou transitória de ácidos nucleicos em células de planta, apenas uma minoria das células absorve o DNA estranho e, se desejado, integra-o em seu genoma, dependendo do vetor de expressão usado e a técnica de transfecção usada. Para identificar e selecionar estes integrantes, um gene que codifica quanto a um marcador selecionável (tal como os descritos acima) é usualmente introduzido nas células hospedeiras junto com o gene de interesse. Estes marcadores pode, por exemplo, ser usados em mutantes em que estes genes são não funcionais, por exemplo, pela anulação por métodos convencionais. Além disso, as moléculas de ácido nucleico que codificam um marcador selecionável podem ser

introduzidas em uma célula hospedeira no mesmo vetor que compreende a sequência que codifica os polipeptídeos a invenção ou usados nos métodos da invenção ou ainda em um vetor separado. As células que foram estavelmente transfectadas com o ácido nucleico introduzido podem ser identificadas, por exemplo, por seleção (por exemplo, as células que integraram o marcador selecionável sobrevivem, visto que as outras células morrem).

Visto que os genes marcadores, particularmente, genes para a resistência a antibióticos e herbicidas, não são mais requeridos ou são indesejados na célula hospedeira transgênica uma vez que os ácidos nucleicos foram introduzidos de maneira bem sucedida, o processo de acordo com a invenção para a introdução dos ácidos nucleicos vantajosamente utilizam técnicas que permitem a remoção ou excisão destes genes marcadores. Um tal método é aquele que é conhecido como co-transformação. O método de co-transformação utiliza dois vetores de maneira simultânea par a transformação, uma vez que o vetor carrega o ácido nucleico de acordo com a invenção e uma segunda safra dos genes marcadores. Uma porção grande de transformantes recebe ou, no caso de plantas, compreende (até 40 % ou mais dos transformantes) ambos os vetores. No caso de transformação com Agrobacteria, os transformantes usualmente recebem apenas uma parte do vetor, isto é, a sequência flanqueada pelo T-DNA, que usualmente representa o cassete de expressão. Os genes marcadores, podem ser subsequentemente removidos a partir da planta transformada realizando-se cruzamentos. Em um outro método, os genes marcadores integrados em um transposon são usados para a transformação junto com o ácido nucleico desejado (conhecido como tecnologia Ac/Ds). Os transformantes podem ser cruzados com uma fonte de transposase ou os transformantes são transformados com uma construção de ácido nucleico que confere a expressão de uma transposase, de maneira transitória ou estável. Em alguns casos (aproximadamente 10 %), o transposon sai do genoma da célula hospedeira uma vez que a transformação

aconteceu de maneira bem sucedida e é perdida. Em diversos outros casos, o transposon vai para um local diferente. Nestes casos o gene marcador deve ser eliminado realizando-se cruzamentos. Na microbiologia, as técnicas foram desenvolvidas tornando-se possível ou facilitando-se a detecção de tais eventos. Um método vantajoso adicionai conta com o que é conhecido como sistemas de recombinação; cuja vantagem é que a eliminação pelo cruzamento pode ser dispensada. O sistema melhor conhecido deste tipo é aquele que é conhecido como o sistema Cre/lox. A Crel é uma recombinase que remove as sequências localizadas entre as sequências loxP. Se o gene marcador for integrado entre as sequências loxP, este é removido uma vez que a transformação aconteceu de maneira bem sucedida, pela expressão da recombinase. Outros sistemas de recombinação adicionais são os sistemas HIN/HIX, FLP/FRT e REP/STB system (Tribble et al., J. Biol. Chem., 275, 2000: 22255-22267; Velmurugan et al., J. Cell Biol., 149, 2000: 553-566). Uma integração específica de local no genoma vegetal das sequências de ácido nucleico de acordo com a invenção é possível. Naturalmente, estes métodos também podem ser aplicados a microorganisms tais como levedura, fungos ou bactérias.

Transgênico/Transgene/Recombinante

Para os propósitos da invenção, "transgênico", "transgene" ou "recombinante" significa com respeito a, por exemplo, uma sequência de ácido nucleico, um cassete de expressão, construção de gene ou um vetor que compreende a sequência de ácido nucleico ou um organismo transformado com as sequências de ácido nucleico, cassetes ou vetores de expressão de acordo com a invenção, todas aquelas construções realizadas pelos métodos recombinantes em que

(a) as sequências de ácido nucleico que codificam proteínas úteis nos métodos da invenção ou

(b) sequências de controle genético que estão

operacionalmente ligadas com a sequência de ácido nucleico de acordo com a invenção, por exemplo um promotor ou

(c) a) e b)

não estão localizados em seu ambiente genético natural ou foram modificados por métodos recombinantes, sendo possível para a modificação tomar a forma de, por exemplo, uma substituição, adição, anulação, inversão ou inserção de um ou mais resíduos de nucleotídeo. O ambiente genético natural é entendido como significando o local genômico ou cromossômico natural na planta original ou na presença em uma biblioteca genômica. No caso de uma biblioteca genômica, o ambiente genético natural da sequência de ácido nucleico é preferivelmente retido, pelo menos em parte. O ambiente flanqueia a sequência de ácido nucleico pelo menos de um lado e tem um comprimento de sequência de pelo menos 50 bp, preferivelmente pelo menos 500 bp, especialmente preferível de pelo menos 1000 bp, mais preferivelmente pelo menos 5000 bp. Um cassete de expressão de ocorrência natural - por exemplo, a combinação de ocorrência natural do promotor natural das sequências de ácido nucleico com a sequência de ácido ácido nucleico correspondente que codifica um polipeptídeo útil nos métodos da presente invenção, como definido acima – torna-se um cassete de expressão transgênico quando este cassete de expressão é modificado pelos métodos não naturais sintéticos ("artificiais"), tais como, por exemplo, tratamento mutagênico. Os métodos adequados são descritos, por exemplo, no US 5.565.350 ou WO 00/15815.

A planta transgênica para os propósitos da invenção é, desta maneira entendida como significando, como acima, que os ácidos nucleicos usados nos métodos da invenção não estão em seu local no genoma da dita planta, sendo possível para os ácidos nucleicos serem expressados de maneira homóloga ou heteróloga. Entretanto, como mencionado, transgênico também significa que enquanto os ácidos nucleicos de acordo com a invenção ou

usados nos métodos inventivos estão em sua posição natural no genoma de uma planta, a sequência foi modificada com respeito à sequência natural e/ou que as sequências reguladoras das sequências naturais foram modificadas. Transgênico é preferivelmente entendido como significando a expressão dos ácidos nucleicos de acordo com a invenção em um local não natural no genoma, isto é, a expressão homóloga ou preferivelmente, heteróloga dos ácidos nucleicos acontece. As plantas transgênicas preferidas são mencionadas neste.

Transformação

O termo "introdução" ou "transformação" como referido neste abrange a transferência de um polinucleotídeo exógeno em uma célula hospedeira, sem consideração ao méodo usado para a transferência. O tecido vegetal capaz de propagação clonal subsequente, por organogênese ou embriogênese, pode ser transformado com uma construção genética da presente invenção e uma planta total regenerada a partir deste. O tecido particular escolhido variará dependendo dos sistemas de propagação clonal disponíveis e melhor adaptadas às espécies particulares sendo transformadas. Os alvos de tecido exemplares incluem discos de folhas, pólen, embriões, cotilédones, hipocotilas, megagametófitos, tecido de calo, tecido meristemático existente (por exemplo, meristema apical, botões axilares e meristemas de raiz) e tecido de meristema induzido (por exemplo, meristema de cotiledônea e meristema de hipocotila). O polinucleotídeo pode ser transitória ou estavelmente introduzido em uma célula hospedeira e pode ser mantido não integrado, por exemplo, como um plasmídeo.

Alternativamente, este pode ser itnegrado no genoma hospedeiro. A célula de planta transformada resultante pode ser então usada para a regeneração de uma planta transformada de uma maneira conhecida pelas pessoas habilitadas na técnica.

A transferência de genes estranhos no genoma de uma planta é

denominada transformação. A transformação de espécies de plantas é agora uma técnica de rotina.vantajosamente, qualquer um de diversos métodos de transformação podem ser usados para a introdução do gene de interesse em uma célula ancestral adequada. Os métodos descritos para a transformação ou regeneração de plantas a partir de tecidos vegetais ou células de planta podem ser utilizados para a transformação estável ou transitória. Os métodos de transformação incluem o uso de lipossomas, eletroporação, produtos químicos que aumentam a entrada de DNA livre, injeção do DNA diretamente na planta, bombardeio de pistola de partículas, transformação usando-se vírus ou pólen e microprojeção. Os métodos podem ser selecionados do método de cálcio/polietileno glicol para protoplastos (Krens, F. A. et al., (1982) Nature 296, 72-74; Negrutiu I et at. (1987) Plant Mol Biol 8: 363-373); eletroporação de protoplastos (Shillito R.D. et al. (1985) Bio/Technol 3, 1099-1102); microinjeção em material vegetal (Crossway A et al., (1986) Mol. Gen Genet 202: 179-185); bombardeio de partícula revestido por DNA ou RNA (Klein TM et al., (1987) Nature 327: 70) infecção com vírus (non-integrativos) e outros. As plantas transgênicas, incluindo plantas de cultivo transgênicas, são preferivelmente produzidas por intermédio da transformação mediada por Agrobacterium. Um método de transformação vantajoso é a transformação *in planta*. Para este fim, é possível, por exemplo, permitir que as agrobactérias atuem em sementes de plantas ou para inocular o meristema da planta com as agrobactérias. Foi mostrado particularmente conveniente de acordo com a invenção permitir que uma suspensão agrobactérias transformadas atuem na planta intacta ou pelo menos nos primórdios do florescimento. A planta é subsequentemente desenvolvidas até as sementes da planta tratada serem obtidas (Clough and Bent, Plant J. (1998) 16, 735-743). Os métodos para a transformação mediada por Agrobacterium do arroz incluem os métodos bem conhecidos para a transformação do arroz, tais como aqueles descritos em qualquer um dos seguintes: Pedido de patente europeu EP 1198985 A1,

Aldemita and Hodges (Planta 199: 612-617, 1996); Chan et al. (Plant Mol Biol 22 (3): 491-506, 1993), Hiei et at. (Plant J 6 (2): 271-282, 1994), cujas divulgações são incorporadas neste por referência como se totalmente apresentado. No caso da transformação do milho, o método preferido é como descrito em Ishida et al. (Nat. Biotechnol 14(6): 745-50, 1996) ou Frame et al. (Plant Physiol 129(1): 13-22, 2002), cujas divulgações são incorporadas por referência neste como se totalmente apresentado. Os ditos métodos ainda são descritos por meio de exemplo em B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S.D. Kung and R. Wu, Academic Press (1993) 128-143 and in Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205-225). Os ácidos nucleicos ou a construção a ser expressada são preferivelmente clonados em um vetor, que é adequado para a trnsformação de *Agrobacterium tumefaciens,* por exemplo, pBin19 (Bevan et at., Nucl. Acids Res. 12 (1984) 8711). As agrobactérias transformadas por um tal vetor podem ser então usadas de maneira conhecida para a transformação de plantas, tais como plantas usadas como um modelo, como *Arabidopsis (Arabidopsis thaliana* está dentro do escopo da presente invenção não considerado como uma planta de cultivo) ou plantas de cultivo, tal como, por meio de exemplo, plantas de tabaco, por exemplo, pelas folhas danificadas por imersão ou folhas cortadas em uma solução agrobacteriana e então cultivando-as em meio adequado. A transformação de plantas por meio de *Agrobacterium tumefaciens* é descrita, por exemplo, por Hofgen and Willmitzer in Nucl. Acid Res. (1988) 16, 9877 ou é conhecido inter alia de F. F. White, Vectors for Gene Transfer in Higher Plants; em Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S.D. Kung and R. Wu, Academic Press, 1993, pp. 15-38.

Além da transformação de células somáticas, quando então devem ser regeneradas em planta sintactas, também é possível transformar as células de meristemas vegetais e, em particular, aquelas células que se

desenvolvem em gametas. Neste caso, os gametas transformados seguem o desenvolvimento natural da planta, dado origem a plantas transgênicas. Desta maneira, por exemplo, as sementes de *Arabidopsis* são tratadas com agrobactérias e as sementes são obtidas a partir do desenvolvimento de plantas que em certa proporção é transformada e, desta maneira, transgênica [Feldman, KA and Marks MD (1987). Mol Gen Genet 208:274-289; Feldmann K (1992). In: C Koncz, N-H Chua and J Shell, eds, Methods in Arabidopsis Research. Word Scientific, Singapore, pp. 274-289]. Os métodos alternativos são fundamentados na remoção repetida das inflorescências e incubação do local de corte no centro da roseta com as agrobactérias transformadas, desse modo, as sementes transformadas também podem ser obtidas em um ponto posterior no tempo (Chang (1994). Plant J. 5: 551-558; Katavic (1994). Mol Gen Genet, 245: 363-370). Entretanto, um método especialmente eficaz é o método de infiltração a vácuo com suas modificações tais como o método de "imersão floral". No caso de infiltração a vácuo de *Arabidopsis,* as plantas intactas sob pressão reduzida são tratadas com uma suspensão agrobacteriana [Bechthold, N (1993). C R Acad Sci Paris Life Sci, 316: 1194-1199], enquanto no caso do método de "imersão floral", o desenvolvimento de tecido floral é incubado brevemente com uma suspensão bacteriana tratada com tensoativo [Clough, SJ and Bent AF (1998) The Plant J. 16, 735-743]. Uma certa proporção de sementes transgênicas é coletada em ambos os casos e estas sementes podem ser distinguidas a partir de sementes não transgênicas desenvolvendo-as sob as condições seletivas descritas acima. além da transformação estável de plastídeos é de vantagem porque os plastídeos são maternalmente herdados e a maioria das lavouras reduzindo ou eliminando o risco do fluxo de transgene através do pólen. A transformação do genoma de cloroplasto é, no geral, atingido por um processo que foi esquematicamente apresentado em Klaus et al., 2004 [Nature Biotechnology 22 (2), 225-229]. Resumidamente, as sequências a serem transformadas são

clonadas junto com um gene de marcador selecionável entre as sequências de flanqueamento homólogas ao genoma de cloroplasto. Estas sequências flanqueadoras homólogas direcionam a integração específica de local no plastoma.

A transformação plastidal foi descrita para muitas espécies vegetais diferentes e uma visão geral é dada em Bock (2001) Transgenic plastids in basic research and plant biotechnology. J Mol Biol. 2001 Sep 21; 312 (3):425-38 ou Maliga, P (2003) Progress towards commercialization of plastid transformação technology. Trends Biotechnol. 21, 20-28. Progresso biotecnológico adicional foi recentemente relatado na forma de transformantes de plastídeo de livre de marcador, que pode ser produzido por um gene marcador co-integrado transitório (Klaus et al., 2004, Nature Biotechnology 22(2), 225-229).

Rotulação de ativação de DNA

A rotulação de ativação de T-DNA (Hayashi et al. Science (1992) 1350-1353), envolve a inserção de T-DNA, usualmente cotnendo um promotor (também pode ser um intensificador de tradução ou um), na região genômica do gene de interesse ou 10 kb até ou a jusante da região codificadora de um gene em uma configuração tal que o promotor direcione a expressão do gene alvejado. Tipicamente, a regulação da expressão do gene alvejado por seu promotor natural é interrompida e o gene cai sob o controle do promotor novamente introduzido. O promotor está tipicamente embutido em um T-DNA. Este T-DNA está aleatoriamente inserido no genoma da planta, por exemplo, através da infecção com *Agrobacterium* e leva à expressão modificada de genes próximos ao T-DNA inserido. As plantas transgênicas resultantes mostram fenotipos dominantes devido à expressão modificada de genes próximos ao promotor introduzido.

TILLING

O termo "TILLING" é uma abreviação de "Lesões Locais

Induzidas Alvejadas nos Genomas" e refere-se a uma tecnologia de mutagênese útil para a geração e/ou identificação de ácidos nucleicos que codificam proteínas com a expressão e/ou a atividade modificada. TILLING também permite a seleção de plantas que carregam tais variantes mutantes. Estas variantes mutantes podem apresentar expressão modificada, em comprimento ou em localização ou em tempo (se as mutações afetarem o promotor, por exemplo). Estas variantes mutantes podem apresentar atividade mais alta do que aquela apresentada pelo gene em sua forma natural. TILLING combina mutagênese de densidade alta com métodos de avaliação de resposta alta. As etapas tipicamente seguindas em TILLING são: (a) mutagênese por EMS (Redei GP and Koncz C (1992) In Methods in Arabidopsis Research, Koncz C, Chua NH, Schell J, eds. Singapore, World Scientific Publishing Co, pp. 16-82; Feldmann et al., (1994) In Meyerowitz EM, Somerville CR, eds, Arabidopsis. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 137-172; Lightner J and Caspar T (1998) In J Martinez-Zapater, J Salinas, eds, Methods on Molecular Biology, Vol. 82. Humana Press, Totowa, NJ, pp 91-104); (b) preparação de DNA e união de indivíduos; (c) amplificação de PCR de uma região de interesse; (d) desnaturação e recozimento para permitir a formação de heteroduplexes; (e) DHPLC, onde a presença de um heteroduplex em um grupo é detectada como um pico extra no cromatograma; (f) identificação do indivíduo mutante e (g) sequenciamento do produto de PCR mutante. Os métodos para TILLING são bem conhecidos na técnica (McCallum et al., (2000) Nat Biotechnol 18: 455-457; revisto por Stemple (2004) Nat Rev Genet 5(2): 145-50).

Recombinação Homóloga

A recombinação homólogapermite a introdução em um genoma de um ácido nucleico selecionado em uma posição selecionada definida. A recombinação homóloga é uma tecnologia padrão usada de maneira rotineira em ciências biológicas para organismos inferiores, tais

como levedura ou o musgo *Physcomitrella.* Os métodos para realizar a recombinação homóloga em plantas foi descrita não apenas para plantas modelos (Offringa et al. (1990) EMBO J 9(10): 3077-84) mas também para plantas de cultivo, por exemplo, arroz (Terada et al. (2002) Nat Biotech 20(10): 1030-4; lida and Terada (2004) Curr Opin Biotech 15(2): 132-8) e os métodos que existem são, no geral, aplicáveis com respeito ao organismo alvo (Miller et al, Nature Biotechnol. 25, 778-785, 2007).

Rendimento

O termo "rendimento", no geral, significa uma produção mensurável de valor econômico, tipicamente relacioando com uma lavoura, a uma área e a um período de tempo especificados. As partes vegetais individuais contribuem diretamente com o rendimento com base em seu número, tamanho e/ou peso ou o rendimento real é o rendimento por metro quadrado para uma lavoura e ano, que é determinadodividindo-se a produção total (inclui tanto produção coletada quanto estimada) por metros quadrados plantados. O termo "rendimento" de uma planta pode se relacionar com a massa vegetativa (raiz e/ou biomassa de broto), para órgãos reprodutivos e/ou aos propagules (tal como sementes) daquela planta.

Vigor precoce

"Vigor precoce" refere-se ao desenvolvimento bem equilibrado saudável ativo durante os estágios precoces do desenvolvimento da planta e pode resultar da aptidão vegetal aumentada devido, por exemplo, às plantas serem melhores adapatadas a seu ambiente (isto é, optimização do uso de recursos de energia e divisão entre broto e raiz). As plantas tendo vigor precoce também apresenta sobrevivência de muda aumentada e um melhor estabelecimento da lavoura, que frequentemente resulta em campos altamente uniformes (com a lavoura desenvolvendo-se de uma maneira uniforme, isto é, com a maioria das plantas atingindo os vários estágiosde desenvolvimento substancialmente ao mesmo tempo) e frequentemente rendimento melhor e

mais alto. Portanto, o vigor precoce pode ser determinado medindo-se vários fatores, tais como peso de semente milhar, porcentagem de germinação, porcentagem de emergência, desenvolvimento de muda, altura da muda, comprimento da raiz, biomassa de broto e de raiz e muito mais.

Aumento/Melhora/Intensificação

Os termos "aumento", "melhora" or "intensificação" são intercambeáveis e devem significar no sentido da aplicação pelo menos a 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 % ou 10 %, preferivelmente pelo menos 15 % ou 20 %, mais preferivelmente 25 %, 30 %, 35 % ou 40 % mais rendimento e/ou desenvolvimento em comparação com as plantas de controle como definido aqui.

Rendimento de semente

O rendimento aumentado de semente pode manifestar-se por si só como um ou mais do seguinte: a) um aumento na biomassa de semente (peso de semente total) que pode estar em uma base de semente individual e/ou por planta e/ou por metro quadrado; b) número aumentado de flores por planta; c) número aumentado de sementes (enchidas); d) taxa de enchimento de semente aumentada (que é expressada como a razão entre o número de sementes enchidas dividido pelo número total de sementes); e) índice de colheita aumentado, que é expressado como a taxa do rendimento de partes colhíveis, tais como sementes, dividido pela biomasa total e f) peso de semente milhar aumentado (TKW), que é extrapolado a partir do número de sementes enchidas contadas em seu peso total. Um TKW aumentado pode resultar de um tamanho de semente aumentado e/ou peso de semente, e pode resultar de um aumento no tamanho de embrião e/ou endosperma.

Um aumento in rendimento de semente também pode ser manifestado como um aumento no tamanho de semente e/ou volume de semente. Além disso, um aumento no rendimento de semente também pode manifestar-se por si só como um aumento na área de semente e/ou

comprimento de semente e/ou largura de semente e/ou perímetro de semente. O rendimento aumentado também pode resultar na arquitetura modificada ou pode ocorrer por causa da arquitetura modificada.

Índice de Verdor

O "índice de verdor" como usado neste é calculado a partir de imagens imagens digitais de plantas. Para cada pixel que pertence ao objeto vegetal na imagem, a razão do valor verde versus o valor vermelho (no modelo RGB para codificar a cor) é calculado. O índice de verdor é expressado como a porcentagem de pixels para os quais a razão verde-para-vermelho excede um dado limiar. Sob condições de de desenvolvimento normais, sob condições de desenvolvimento com tensão por sal e sob condições de desenvolvimento com disponibilidade de nutriente reduzida, o índice de verdor das plantas é medido na última formação de imagem antes do florescimento. Em contraste, sob condições condições de desenvolvimento com tensão por secura, o índice de verdor das plantas é medido na primeira formação de imagem após a secura.

Planta

O termo "planta" como usado neste abrange plantas totais, ancestrais e progênie das plantas e partes de plantas, incluindo sementes, brotos, caules, folhas, raízes (incluindo tubérculos), flores e tecidos e órgãos, em que cada um dos já mencionados compreendem o gene/ácido nucleico de interesse. O termo "planta" também abrange células de planta, culturas de suspensão, tecido de calo, embriões, regiões meristemáticas, gametófitos, esporófitos, pólen e microesporos, again em que cada um dos já mencionados compreendem o gene/ácido nucleico de interesse.

As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular, plantas monocotiledôneas e dicotiledôneas incluindo forragem ou legumes de forragem, plantas ornamentais, lavouras de

alimentos, árvores ou arbustos selecionados da lista que compreende *Acer* spp., *Actinidia* spp., *Abelmoschus* spp., *Agave sisalana, Agropyron* spp., *Agrostis stolonifera, Allium* spp., *Amaranthus* spp., *Ammophila arenaria, Ananas comosus, Annona* spp., *Apium graveolens, Arachis spp, Artocarpus* spp., *Asparagus officinalis, Avena* spp. (por exemplo, *Avena sativa, Avena fatua, Avena byzantina, Avena fatua var. sativa, Avena hybrida), Averrhoa carambola, Bambusa sp., Benincasa hispida, Bertholletia excelsea, Beta vulgaris, Brassica* spp. (por exemplo, *Brassica napus, Brassica rapa* ssp. [canola, óleo de semente de colza, nabo]), *Cadaba farinosa, Camellia sinensis, Canna indica, Cannabis sativa, Capsicum* spp., *Carex a/ata, Carica papaya, Carissa macrocarpa, Carya* spp., *Carthamus tinctorius, Castanea* spp., *Ceiba pentandra, Cichorium endivia, Cinnamomum* spp., *Citrullus lanatus, Citrus* spp., *Cocos* spp., *Coffea* spp., *Colocasia esculenta, Cola* spp., *Corchorus sp., Coriandrum sativum, Corylus spp., Crataegus* spp., *Crocus sativus, Cucurbita* spp., *Cucumis spp., Cynara* spp., *Daucus carota, Desmodium spp., Dimocarpus longan, Dioscorea* spp., *Diospyros* spp., *Echinochloa* spp., *Elaeis (por exemplo, Elaeis guineensis, Elaeis oleifera), Eleusine coracana, Eragrostis tef, Erianthus sp., Eriobotrya japonica, Eucalyptus sp., Eugenia uniflora, Fagopyrum* spp., *Fagus* spp., *Festuca arundinacea, Ficus carica, Fortunella* spp., *Fragaria* spp., *Ginkgo biloba, Glycine* spp. (por exemplo, *Glycine max, Soja hispida* ou *Soja max), Gossypium hirsutum, Helianthus spp. (por exemplo, Helianthus annuus), Hemerocallis fulva, Hibiscus* spp., *Hordeum* spp. (por exemplo, *Hordeum vulgare), lpomoea batatas, Juglans* spp., *Lactuca sativa, Lathyrus spp., Lens culinaris, Linum usitatissimum, Litchi chinensis, Lotus spp., Luffa acutangula, Lupinus* spp., *Luzula sylvatica, Lycopersicon* spp. (por exemplo, *Lycopersicon esculentum, Lycopersicon lycopersicum, Lycopersicon pyriforme), Macrotyloma* spp., *Malus* spp., *Malpighia emarginata, Mammea americana, Mangifera indica, Manihot* spp., *Manilkara zapota, Medicago*

sativa, Melilotus spp., *Mentha* spp., *Miscanthus sinensis, Momordica* spp., *Morus nigra, Musa* spp., *Nicotiana* spp., *Olea* spp., *Opuntia spp., Omithopus spp., Oryza spp. (por exemplo, Oryza sativa, Oryza latifolia), Panicum miliaceum, Panicum virgatum, Passiflora edulis, Pastinaca sativa, Pennisetum sp., Persea spp., Petroselinum crispum, Phalaris arundinacea, Phaseolus* spp., *Phleum pratense, Phoenix* spp., *Phragmites australis, Physalis* spp., *Pinus spp., Pistacia vera, Pisum* spp., *Poa* spp., *Populus* spp., *Prosopis spp., Prunus spp., Psidium* spp., *Punica granatum, Pyrus communis, Quercus* spp., *Raphanus sativus, Rheum rhabarbarum, Ribes* spp., *Ricinus communis, Rubus spp., Saccharum* spp., *Salix sp., Sambucus spp., Secale cereale, Sesamum* spp., *Sinapis sp., Solanum spp. (por exemplo, Solanum tuberosum, Solanum integrifolium* ou *Solanum lycopersicum), Sorghum bicolor, Spinacia* spp., *Syzygium* spp., Tagetes spp., *Tamarindus indica, Theobroma cacao, Trifolium* spp., *Tripsacum dactyloides, Triticosecale rimpaui, Triticum* spp. (por exemplo, *Triticum aestivum, Triticum durum, Triticum turgidum, Triticum hybernum, Triticum* macha, *Triticum sativum, Triticum monococcum ou Triticum vulgare), Tropaeolum minus, Tropaeolum majus, Vaccinium* spp., *Vicia* spp., *Vigna* spp., *Viola* odorata, *Vitis* spp., Zea *mays, Zizania palustris, Ziziphus* spp., entre outros.

Descrição Detalhada da Invenção

I. Proteína que compreende o domínio LOB (LOB: Fronteiras de Órgão Lateral) Surpreendentemente, foi observado agora que modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo LBD dá plantas tendo traços intensificados relacionados com o rendimento com relação às plantas de controle. De acordo com uma primeira forma de realização, a presente invenção fornece um método para intensificar traços relacionados com o rendimento em plantas com relação às plantas de controle, que compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo LBD.

Um método preferido para modular (preferivelmente, aumentar) a expressão de um ácido nucleico que codifica um polipeptídeo LBD é pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo LBD.

Qualquer referência a seguir a uma "proteína útil nos métodos da invenção" é determinada entender um polipeptídeo LBD como definido neste. Qualquer referência a seguir a um "ácido nucleico útil nos métodos da invenção" é determinada entender um ácido nucleico capaz de codificar um tal polipeptídeo LBD. O ácido nucleico a ser introduzido em uma planta (e, portanto, útil na realização dos métodos da invenção) é qualquer ácido nucleico que codifica o tipo de proteína que será descrito agora, a seguir também denominado "ácido nucleico LBD" ou "gene LBD".

A "polipeptídeo LBD" como definido neste refere-se a qualquer polipeptídeo que compreenda um domínio DUF260 (número de acessão Pfam PF03195, número de acessão Interpro IPR004883, ver Figura 1). Preferivelmente, a sequência de polipeptídeo LBD, quando usada na construção de uma árvore filogenética tal como a descrita na Figura 3 (ou como definido por Yang et al., 2006, usando-se o método HKY (Hasegawa et al., J. Mol. Evol. 22, 160-174, 1985)), agrupamentos com o grupo de polipeptídeos de domínio LOB de classe II que compreende a sequência de aminoácido representada pela SEQ ID N°: 2 em vez de qualquer outro grupo.

Ainda preferivelmente, a proteína LBD compreende pelo menos um dos seguintes motivos conservados:

Motivo 1: MSCNGCRXLRKGCX (SEQ ID N°: 5); em que X na posição 8 pode ser qualquer aminoácido, mas, preferivelmente V ou I e em que X na posição 14 pode ser qualquer aminoácido, mas, preferivelmente é um de S, G ou N.

Motivo 2: QXXATXFXAKFXGR (SEQ ID N°: 6), em que X na posição 2 pode ser qualquer aminoácido, mas, preferivelmente um de A, S,

ou G; em que X na posição 3 pode ser qualquer aminoácido, mas, preferivelmente um de N, Q ou H; em que X na posição 6 pode ser qualquer aminoácido, mas, preferivelmente um de V, L ou I; em que X na posição 8 pode ser qualquer aminoácido, mas, preferivelmente um de L, V, A ou I e em que X na posição 12 pode ser qualquer aminoácido, mas, preferivelmente um de Y ou F.

Motivo 3: FXSLLXEAXG (SEQ ID N°: 7); em que X na posição 2 pode ser qualquer aminoácido, mas, preferivelmente um de R, S, K ou Q; em que X na posição 6 pode ser qualquer aminoácido, mas, preferivelmente um de Y, H ou F e em que X na posição 9 pode ser qualquer aminoácido, mas, preferivelmente C ou A.

Ainda preferivelmente, a proteína LBD compreende mais do que sete, mais preferivelmente pelo menos do que nove resíduos Cys e não compreende a assinatura DP(V/I)YG (SEQ ID N°: 8).

O domínio DUF260 é caracterizado pela presença de um motivo C-x(2)-C-x(6)-C-x(3)-C (SEQ ID N°: 9), em que X pode ser qualquer aminoácido.

O termo "domínio" e "motivo é definido na seção "definições" aqui. Bancos de dados especialistas existem para a identificação de domínios, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-244, InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318, Prosite (Bucher and Bairoch (1994), uma sintaxe de perfil generalizado para motivos de sequências biomoleculares e sua função na interpretação de sequência automática. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004) ou Pfam (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002). Uma série de ferramentas para a

análise *in silico* de sequências de proteínas está disponível no servidor ExPASy proteomics (Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31:3784-3788(2003)). Os domínios também são identificados usando-se técnicas de rotina, tal como pelo alinhamento de sequência.

Os métodos para o alinhamento de sequências para comparação são bem conhecidos na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. GAP usa o algoritmo de Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar o alinhamento global (isto é, medindo-se as sequências completas) de duas sequências que maximizam o número de equiparações e minimiza o número de fendas. O algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-10) calcula a identidade de sequência percentual e realiza uma análise estatística da similaridade entre as duas sequências. O software para realizar a análise BLAST está publicamente disponível através do National Centre for Biotechnology Information (NCBI). os homólogos podem ser facilmente identificados usando-se, por exemplo, o algoritmo de alinhamento de sequência múltipla ClustalW (versão 1.83), com os parâmetros de alinhamento default e um método de registro em porcentagem. As porcentagens globais de similaridade e de identidade também podem ser determinados usando-se um dos métodos disponíveis no pacote de software MatGAT (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando-se proteína ou sequências de DNA). A edição manual menor pode ser realizada para otimizar o alinhamento entre os motivos conservados, como deve estar evidente para uma pessoa habilitada na técnica. Além disso, em vez de usar sequências de comprimento total para a identificação de homólogos, domínios específicos também podem ser usados. Os valores de identidade de

sequência podem ser determinados sobre o aminoácido inteiro ou sequência de aminoácido ou em domínios selecionados ou motivos conservados, usando-se os programas mencionados acima usando-se os parâmetros default.

Além disso, os polipeptídeos LBD (pelo menos em sua forma natural), como fatores de transcrição, tipicamente têm atividade de ligação de DNA. Ferramentas e técnicas para medir a atividade de ligação de DNA são bem conhecidas na técnica e incluem por exemplo ensaios de retardo de gel. Os métodos experimentais para caracterizar a atividade fatores de transcrição podem ser encontrados, por exemplo, em Sambrook et al. (2001) Molecular Cloning: a laboratory manual, 3° Edição Cold Spring Harbor Laboratory Press, CSH, New York ou nos Volumes 1 e 2 de Ausubel et al. (atualizado anualmente), Current Protocols in Molecular Biology, Current Protocols.

A presente invenção é ilustrada transformando-se as plantas com a sequência de ácido nucleico representada pela SEQ ID N°: 1, que codifica a sequência de polipeptídeo da SEQ ID N°: 2. Entretanto, o desempenho da invenção não está restrito a estas sequências; os métodos da invenção podem ser vantajosamente realizados usando-se qualquer ácido nucleico que codifica LBD ou polipeptídeo LBD como definido neste.

Os exemplos de ácidos nucleicos que codificam os polipeptídeos LBD são dadas na Tabela A1 do Exemplo 1 neste. Tais ácidos nucleicos são úteis na realização dos métodos da invenção. As sequências de aminoácido dadas na Tabela A1 do Exemplo 1 são sequências de exemplos de ortólogos e parálogos do polipeptídeo LBD representado pela SEQ ID N°: 2, os termos "ortólogos" e "parálogos" sendo como definido neste. Ortólogos e parálogos adicionais podem ser facilmente definidos pela realização de uma denominada pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envovle submeter uma sequência em dúvida ao BLAST (por exemplo, usando-se qualquer uma das sequências listadas na Tabela A1 do Exemplo 1) contra qualquer banco de dados de sequência, tais como os

bancos de dados NCBI publicamente disponíveis. BLASTN ou TBLASTX (usando-se os valores default padrão) são, no geral, usados quando começa-se a partir de uma sequência de nucleotídeo e BLASTP ou TBLASTN (usando-se os valores default padrão) quando começa-se a partir de uma sequência de proteína. Os resultados de BLAST podem ser opcionalmente filtrados. As sequências de comprimento total de cada um dos resultados filtrados ou resultados não filtrados são então submetidos ao BLAST novamente (segundo BLAST) contra sequências do organismo a partir do qual a sequência em questão é derivada (quando a sequência em questão é a SEQ ID N°: 1 ou SEQ ID N°: 2, o segundo BLAST portanto, deve ser contra sequências de Arabidopsis). os resultados do primeiro e segundo BLASTs são então comparados. um parálogo é identificado se um acerto de classificação alta do primeiro blast for da mesma espécie como a partir da qual a sequência em questão é derivada, submetendo-se ao BLAST novamente então resulta idealmente na sequência em questão entre os acertos mais altos; um ortólogo é identificado se um acerto de classificação alta no primeiro BLAST não for da mesma espécie como a partir da qual a sequência em questão é derivada e preferivelmente resulta no BLAST novamente na sequência em questão estando entre os acertos mais altos.

Os acertos de classificação alta são aqueles tendo um valor E baixo. Quanto mais baixo o valor E, mais significante o registro (ou, em outras palavras, menor a chance que o acerto seja encontrado por chance). A computação do valor E é bem conhecido na técnica. Além dos valores E, as comparações são registradas pela identidade de porcentagem. A Identidade de porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácidos nucleicos comparadas (ou polipeptídeo) em um comprimento particular. No caso de famílias grandes, o ClustalW pode ser usado, seguido por uma árvore de união próxima, para ajudar a visualizar o agrupamento de genes relacionados e para identificar ortólogos e

parálogos.

As variantes de ácido nucleico também podem ser úteis na prática dos métodos da invenção. Os exemplos de tais variantes incluem ácidos nucleicos que codificam homólogos e derivados de qualquer uma das sequências de aminoácido dadas na Tabela A1 do Exemplo 1, os termos "homólogo" e "derivado" sendo como definido neste. Também são úteis nos métodos da invenção os ácidos nucleicos que codificam homólogos e derivados de ortólogos ou parálogos de qualquer uma das sequências de aminoácido dadas na Tabela A1 do Exemplo 1. Os homólogos e derivados úteis nos métodos da presente invenção têm substancialmente a mesma atividade biológica e funcional como a proteína não modificada da qual estes são derivados.

As Variantes de ácido nucleico adicionais úteis na prática dos métodos da invenção incluem porções de ácidos nucleicos que codificam polipeptídeos LBD, ácidos nucleicos que hibridizam-se a ácidos nucleicos que codificam polipeptídeos LBD, variantes de união de ácidos nucleicos que codificam polipeptídeos LBD, variantes alélicas de ácidos nucleicos que codificam polipeptídeos LBD e variantes de ácidos nucleicos que codificam polipeptídeos LBD obtidos pelo embaralhamento de gene. Os termos sequência de hibridização, variante de união, variante alélica e embaralhamento de gene são como descritos neste.

Vantajosamente, a presente invenção fornece até agora sequências de ácido nucleico LBD e polipeptídeo desconhecidas.

De acordo com uma outra forma de realização da presente invenção, é fornecido uma molécula de ácido nucleico isolada que compreende:

(i) um ácido nucleico representado pela SEQ ID N°: 69;

(ii) um ácido nucleico ou fragmento do mesmo que é complementar a qualquer uma das SEQ ID N° dadas em (i);

(iii) um ácido nucleico que codifica um polipeptídeo LBD tendo, na ordem crescente de preferência, pelo menos 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência a SEQ ID N°: 70;

(iv) um ácido nucleico capaz de hibridizar sob condições estringentes a qualquer um dos ácidos nucleicos dados em (i), (ii) ou (iii) acima.

De acordo com uma outra forma de realização da presente invenção, portanto, é fornecido um polipeptídeo isolado que compreende:

(i) uma sequência de aminoácido tendo, na ordem crescente de preferência, pelo menos 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou 100 % de identidade de sequência à sequência de aminoácido dada na SEQ ID N°: 70.

(ii) derivados de qualquer uma das sequências de aminoácido dadas em (i).

Os ácidos nucleicos que codificam polipeptídeos LBD não necessitam ser ácidos nucleicos de comprimento total, visto que o desempenho dos métodos da invenção não contam com o uso de sequências de comprimento total de ácido nucleico. De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta uma porção de qualquer uma das sequências de ácido nucleico dadas na Tabela A1 do Exemplo 1 ou uma porção de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de aminoácido dadas na Tabela A1 do Exemplo 1.

Uma porção de um ácido nucleico pode ser preparada, por exemplo, realizando-se uma ou mais anulações no ácido nucleico. As porções podem ser usadas na forma isolada ou estas podem ser fundidas a outras sequências codificadoras (ou não codificadoras) a fim de, por exemplo,

produzir uma proteína que combine diversas atividades. Quando fundido a outras sequências codificadoras, o polipeptídeo resultante produzido na tradução pode ser maior do que aquele predito para a porção de proteína.

As porções úteis nos métodos da invenção, codificam um polipeptídeo LBD como definido neste têm substancialmente a mesma atividade biológica como as sequências de aminoácido dadas na Tabela A1 do Exemplo 1. Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos dados na Tabela A1 do Exemplo 1 ou é uma porção de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das sequências de aminoácido dadas na Tabela A1 do Exemplo 1. Preferivelmente a porção é de pelo menos 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 nucleotídeos consecutivos de comprimento, os nucleotídeos consecutivos sendo de qualquer uma das sequências de ácido nucleico dadas na Tabela A1 do Exemplo 1 ou de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das sequências de aminoácido dadas na Tabela A1 do Exemplo 1. Mais preferivelmente, a porção é uma porção do ácido nucleico da SEQ ID N°: 1. Preferivelmente, a porção codifica uma sequência de aminoácido que quando usada na construção de uma árvore filogenética tal como a descrita na Figura 3 (ou como definido por Yang et al., 2006, usando-se o método HKY (Hasegawa et al., J. Mol. Evol. 22, 160-174, 1985)), agrupamentos com o grupo de polipeptídeos de domínio LOB de classe II que compreende a sequência de aminoácido representada pela SEQ ID N°: 2 em vez de qualquer outro grupo.

Uma outra variante de ácido nucleico útil nos métodos da invenção é um ácido nucleico capaz de hibridizar, sob condições de estringência reduzida, preferivelmente sob condições estringentes, com um ácido nucleico que codifica um polipeptídeo LBD como definido neste ou com uma porção como definido neste.

De acordo com a presente invenção, é fornecido um método

para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta, um ácido nucleico capaz de hibridizar qualquer um dos ácidos nucleicos dados na Tabela A1 do Exemplo 1 ou que compreende introduzir e expressar em uma planta, um ácido nucleico capaz de hibridizar a um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de ácido nucleico dadas na Tabela A1 do Exemplo 1.

As sequências de hibridização úteis nos métodos da invenção codificam um polipeptídeo LBD como definido neste e têm substancialmente a mesma atividade biológica como as sequências de aminoácido dadas na Tabela A1 do Exemplo 1. Preferivelmente, a sequência de hibridização é capaz de hibridizar a qualquer um dos ácidos nucleicos dados na Tabela A1 do Exemplo 1 ou a uma porção de qualquer destas sequências, uma porção sendo como definida acima ou em que a sequência de hibridização é capaz de hibridizar a um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das sequências de aminoácido dadas na Tabela A1 do Exemplo 1. Mais preferivelmente, a sequência de hibridização é capaz de hibridizar a um ácido nucleico como representado pela SEQ ID N°: 1 ou a uma porção deste.

Preferivelmente, a sequência de hibridização codifica uma sequência de aminoácido que quando usada na construção de uma árvore filogenética tal como a descrita na Figura 3 (ou como definido por Yang et al., 2006, usando-se o método HKY (Hasegawa et al., J. Mol. Evol. 22, 160-174, 1985)), agrupamentos com o grupo de polipeptídeos de domínio LOB de classe II que compreende a sequência de aminoácido representada pela SEQ ID N°: 2 em vez de qualquer outro grupo.

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante de união que codifica um polipeptídeo LBD como definido neste acima, a variante de união sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta a variante de união de qualquer uma das sequências de ácido nucleico dadas na Tabela A1 do Exemplo 1 ou a variante de união de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de aminoácido dadas na Tabela A1 do Exemplo 1.

As variantes de união preferidas são variantes de união de um ácido nucleico representado pela SEQ ID N°: 1 ou a variante de união de um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 2. Preferivelmente, a sequência de aminoácido codificada pela variante de união, quando usado na construção de uma árvore filogenética tal como a descrita na Figura 3 (ou como definido por Yang et al., 2006, usando-se o método HKY (Hasegawa et al., J. Mol. Evol. 22, 160-174, 1985)), agrupamentos com o grupo de polipeptídeos de domínio LOB de classe II que compreende a sequência de aminoácido representada pela SEQ ID N°: 2 em vez de qualquer outro grupo.

Uma outra variante de ácido nucleico útil na realização dos métodos da invenção é uma variante alélica de um ácido nucleico que codifica um polipeptídeo LBD como definido neste acima, uma variante alélica sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta uma variante alélica de qualquer um dos ácidos nucleicos dados na Tabela A1 do Exemplo 1 ou que compreende introduzir e expressar em uma planta uma variante alélica de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de aminoácido dadas na Tabela A1 do Exemplo 1.

As variantes alélicas úteis nos métodos da presente invenção

têm substancialmente a mesma atividade biológica como o polipeptídeo LBD da SEQ ID N°: 2 qualquer um dos aminoácidos descritos na Tabela A1 do Exemplo 1. As variante alélicas que existem em estado natural e abrangidas dentro dos métodos da presente invenção são o uso destes alelos naturais. Preferivelmente, a variante alélica é uma variante alélica da SEQ ID N°: 1 ou uma variante alélica de um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 2. Preferivelmente, a sequência de aminoácido codificada pela variante alélica, quando usada na construção de uma árvore filogenética tal como a descrita na Figura 3 (ou como definido por Yang et al., 2006, usando-se o método HKY (Hasegawa et al., J. Mol. Evol. 22, 160-174, 1985)), agrupamentos com o grupo de polipeptídeos de domínio LOB de classe II que compreende a sequência de aminoácido representada pela SEQ ID N°: 2 em vez de qualquer outro grupo.

Mistura de gene e a evolução direcionada podem ser usadas para gerar variantes de ácidos nucleicos que codificam polipeptídeos LBD como definido acima; o termo "mistura de gene" sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta uma variante de qualquer uma das sequências de ácido nucleico dadas na Tabela A1 do Exemplo 1 ou que compreende introduzir e expressar em uma planta uma variante de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de aminoácido dadas na Tabela A1 do Exemplo 1, cujo ácido nucleico variante é obtido pelo embaralhamento de gene.

Preferivelmente, a sequência de aminoácido codificada pelo ácido nucleico variante obtido pelo embaralhamento de gene, quando usado na construção de uma árvore filogenética tal como a descrita na Figura 3 (ou como definido por Yang et al., 2006, usando-se o método HKY (Hasegawa et al., J. Mol. Evol. 22, 160-174, 1985)), agrupamentos com o grupo de

polipeptídeos de domínio LOB de classe II que compreende a sequência de aminoácido representada pela SEQ ID N°: 2 em vez de qualquer outro grupo.

Além disso, as variantes de ácido nucleico também podem ser obtidas por mutagênese direcionada ao local. Diversos métodos estão disponíveis para atingir a mutagênese direcionada ao local, o mais comum sendo Métodos com base em PCR (Current Protocols in Molecular Biology. Wiley Eds.).

Os ácidos nucleicos que codificam polipeptídeos LBD podem ser derivados de qualquer fonte natural ou artificial. O ácido nucleico pode ser modificado a partir de sua forma natural na composição e/ou ambiente genômico através da manipulação humana deliberada. Preferivelmente, o ácido nucleico que codifica o polipeptídeo LDB é de uma planta, ainda preferivelmente de uma planta dicotiledônea, mais preferivelmente da família Brassicaceae, mais preferivelmente o ácido nucleico é de *Arabidopsis thaliana.*

O desempenho dos métodos da invenção dá plantas tendo traços intensificados relacionados com o rendimento. Em particular, o desempenho dos métodos da invenção dá plantas tendo rendimento aumentado, especialmente rendimento aumentado de semente com relação às plantas de controle. Os termos "rendimento" e "rendimento de semente" são descritos em mais detalhes na seção "definições" neste.

Referência neste a traços intensificados relacionados com o rendimento é determinada entender um aumento na biomassa (peso) de uma ou mais partes de uma planta, que pode incluir partes (coletáveis) acima do solo e/ou partes (coletáveis) abaixo do solo. Em particular, tais partes colhíveis acima do solo são biomassa de broto e sementes, and desempenho dos métodos da invenção resulta em plantas tendo biomassa aumentada de broto e rendimento de semente com relação à biomassa de broto e rendimento de semente de plantas de controle.

Tomando o milho como um exemplo, um aumento no rendimento pode ser manifestado como um ou mais do seguinte: aumento no número de plantas estabelecido por hectare ou acre, um aumento no número de espigas por planta, um aumento no número de fileiras, número de sementes por fileira, peso de semente, peso de semente milhar, comprimento/diâmetro de espiga, aumento na razão de enchimento de semente (que é o número de sementes enchidas dividido pelo número total de sementes e multiplicado por 100), entre outros. Tomando-se o arroz como um exemplo, um aumento no rendimento pode manifestar-se por si só como um aumento em um ou mais do seguinte: número de plantas por hectare ou acre, número de panículas por planta, número de espigas por panícula, número de flores (flósculos) por panícula (que é expressado como uma razão do número de sementes enchidas no número de panícula primárias), aumento na razão de enchimento de semente (que é o número de sementes enchidas dividido pelo número total de sementes e multiplicado por 100), aumento em peso de semente milhar, entre outros.

A presente invenção fornece um método para aumentar o rendimento, especialmente rendimento de biomassa de broto e rendimento de semente de plantas, com relação às plantas de controle, cujo método compreende a expressão modular, preferivelmente aumentar a expressão, em uma planta de um ácido nucleico que codifica um polipeptídeo LBD como definido neste.

Visto que as plantas transgênicas de acordo com a presente invenção têm rendimento aumentado, é provável que estas plantas apresentem uma taxa de desenvolvimento aumentada (durante pelo menos parte do seu ciclo de vida), com relação à taxa de desenvolvimento de plantas de controle em um estágio correspondente em seu ciclo de vida.

A taxa de desenvolvimento aumentada pode ser específica para uma ou mais partes de uma planta (incluindo sementes) ou pode ser

substancialmente por toda a planta. As plantas tendo uma taxa de desenvolvimento aumentada pode ter um ciclo de vida mais curto. o ciclo de vida de uma planta pode ser tomado significar o tempo necessário para o desenvolvimento de uma semente madura seca até o estágio onde a planta produziu sementes maduras secas, similar ao material de partida. Este ciclo de vida pode ser influenciado por fatores tais como vigor precoce, taxa de desenvolvimento, índice de verdor, tempo de florescimento e velocidade de maturação de semente. O aumento em taxa de desenvolvimento pode acontecer em um ou mais estágios no ciclo de vida de uma planta ou substancialmente durante o ciclo de vida de planta total. A taxa de desenvolvimento aumentada durante os estágios precoces no ciclo de vida de uma planta pode refletir o vigor intensificado. O aumento na taxa de desenvolvimento pode alterar o ciclo de coleta de uma planta permitindo que as plantas sejam semeadas depois e/ou coletadas mais cedo do que seria de outra maneira possível (um efeito similar pode ser obtido com tempo de florescimento precoce). Se a taxa de desenvolvimento for suficientemente aumentada, isto pode permitir a semeadura adicional de sementes das mesmas espécies vegetais (por exemplo, semeadura e coleta de plantas de arroz seguido pela semeadura e coleta de outras plantas de arroz tudo dentro de um período de desenvolvimento convencional). Similarmente, se a taxa de desenvolvimento for suficientemente aumentada, isto pode permitir a semeadura adicional de sementes de espécies de plantas diferentes (por exemplo a semeadura e a coleta de plantas de milho seguido por, por exemplo, a semeadura e a coleta opcional de soja, batata ou qualquer outro vegetal adequado). Os Períodos de coleta adicionais a partir do mesmo estoque de raiz no caso de algumas plantas de cultivo também pode ser possível. A alteração do ciclo de coleta de uma planta pode levar a um aumento na produção de biomassa anual por acre (devido a um aumento no número de vezes (diz-se em um ano) que qualquer planta particular pode ser

desenvolvida e coletada). Um aumento na taxa de desenvolvimento também pode permitir o cultivo de plantas transgênicas em uma área geográfica mais ampla do que suas contrapartes do tipo selvagem, visto que as limitações territoriais para o desenvolvimento de uma lavoura são frequentemente determinados pelas condições ambientais adversas no período de plantio (estação precoce) ou no período de coleta (estação posterior). Tais condições adversas podem ser evitadas se o ciclo de coleta for encurtado. A taxa de desenvolvimento pode ser determinada derivando-se vários parâmetros de curvas de desenvolvimento, tais parâmetros podem ser: T-Mid (o tempo levado pelas plantas para atingir 50 % de seu tamanho máximo) e T-90 (tempo levado pelas plantas para atingir 90 % de seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção dá plantas tendo uma taxa de desenvolvimento aumentada com relação às plantas de controle. Portanto, de acordo com a presente invenção, é fornecido um método for aumentar a taxa de desenvolvimento de plantas, cujo método compreende a expressão modular, preferivelmente aumentar a expressão, em uma planta de um ácido nucleico que codifica um polipeptídeo LBD como definido neste. Em particular, um aumento na taxa de desenvolvimento foi observada durante os estágios de desenvolvimento precoce da planta (vigor precoce).

Um aumento no rendimento e/ou taxa de desenvolvimento ocorre se a planta estiver sob condições sem estresse ou se a planta for exposta a várias tensões em comparação com as plantas de controle. As Plantas tipicamente respondem à exposição à tensão desenvolvendo-se mais lentamente. Em condições de tensão grave, a planta pode ainda parar de se desenvolver completamente. A tensão branda por outro lado é definida neste como sendo qualquer tensão à qual uma planta é exposta que não resulta na planta parando de se desenvolver completamente sem sem a capacidade de

retomar o crescimento. A tensão branda no sentido da invenção da invenção leva a uma redução no desenvolvimento das plantas submetidas à tensão de menos que 40 %, 35 % ou 30 %, preferivelmente menor do que 25 %, 20 % ou 15 %, mais preferivelmente menor do que 14 %, 13 %, 12 %, 11 % ou 10 % ou menor em comparação com a planta de controle sob condições sem estresse. Devido aos avanços nas práticas agrícolas (irrigação, fertilização e tratamentos com pesticidas) tensões graves não são frequentemente encontradas em plantas de cultivo cultivadas. Como uma consequência, o desenvolvimento comprometido induzido pela tensão branda é frequentemente uma característica indesejada para a agricultura. As tensões brandas são as tensões (ambientais) bióticas e/ou abióticas diárias às quais a planta é exposta. As tensões abióticas podem ocorrer devido à secura ou excesso de água, tensão anaeróbica, tensão ao sal, toxicidade química, tensão oxidativa e calor, frio ou temperaturas congelantes. A tensão abiótica pode ser uma tensão osmótica causada por uma tensão a água (particularmente devido à secura), tensão ao sal, tensão oxidativa ou uma tensão iônica. As Tensões bióticas são tipicamente aquelas tensões causadas por patogênio, tais como bactérias, vírus, fungos e insetos.

Em particular, os métodos da presente invenção podem ser realizados sob condições sem estresse ou sob condições de secura branda para dar plantas tendo rendimento aumentado com relação às plantas de controle. Como relatado em Wang et al. (Plant (2003) 218: 1-14), a tensão abiótica leva a uma série de mudanças morfológicas, fisiológica, bioquímicas e moleculares que afetam adversamente o desenvolvimento e produtividade da planta. Secura, salinidade, temperaturas extremas e tensão oxidativa são conhecidas serem interconectadas e podem induzir o desenvolvimento e o dano celular através de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve um grau particularmente alto de "cruzamento" de estresse por secura e tensão por salinidade alta. Por

exemplo, seca e/ou desalinização são manifestadas primariamente como tensão osmótica, resultando na interrupção da homeostase e a distribuição de íon na célula. Tensão oxidativa, que frequentemente acompanha temperatura alta e baixa, tensão por salinidade ou secura, pode causar a desnaturação de de proteínas funcionais e estruturais. Como uma consequência, estas diversas tensões ambientais frequentemente ativam caminhos de sinalização celular e respostas celulares similares, tais como a produção de proteínas de tensão, super-regulação do acúmulo de anti-oxidantes, acúmulo dos solutos compatíveis e interrupção do desenvolvimento. O termo condições "sem tensão" como usado aqui são aquelas condições ambientais que permitem o desenvolvimento ótimo de planta de plantas. As pessoas habilitadas na técnica estão cientes das condições de solo normais e condições climáticas para uma dada localização. As plantas com condições de desenvolvimento ótimas, (desenvolvem-se sob condições sem estresse) tipicamente, o rendimento na ordem crescente de preferência pelo menos 90 %, 87 %, 85 %, 83 %, 80 %, 77 % ou 75 % da produção média de tal planta em um dado ambiente. A produção média pode ser calculada com base na coleta e/ou na estação. As pessoas habilitadas na técnica estão cientes das produções de rendimento médio de uma lavoura.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições sem estresse ou sob condições de secura brandas com rendimento aumentado com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas que se desenvolvem sob condições sem estresse ou sob condições de secura brandas, cujo método compreende aumentar expressão em uma planta de um ácido nucleico que codifica um polipeptídeo LBD.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições de deficiência de nutriente, particularmente sob

condições de deficiência de nitrogênio, vigor precoce aumentado e rendimento aumentado com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método for aumentar vigor precoce e rendimento em plantas desenvolvidas sob condições de deficiência de nutriente, cujo método compreende aumentar a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo LBD. A Deficiência de nutriente pode resultar de uma perda ou de um excesso de nutrientes tais como nitrogênio, fosfatos e outros compostos contendo fósforo, potássio, cálcio, cádmio, magnésio, manganês, ferro e boro, entre outros.

A presente invenção abrange plantas ou partes destas (incluindo sementes) obteníveis pelos métodos de acordo com a presente invenção. As plantas ou partes destas compreende um ácido nucleico transgene que codifica um polipeptídeo LBD como definido acima.

A invenção também fornece contruções genéticas e vetores para facilitar a introdução e/ou expressão em plantas de ácidos nucleicos que codificam polipeptídeos LBD. A construção de genes pode ser inserida em vetores, que pode estar comercialmente disponível, adequada para a transformação em plantas e adequado para a expressão do gene de interesse nas células transformadas. A invenção também forence o uso da construção de gene como definido neste nos métodos da invenção.

Mais especificamente, a presente invenção fornece uma construção que compreende:

(a) um ácido nucleico que codifica um polipeptídeo LBD como definido acima;

(b) uma ou mais sequências de controle capazes de conduzir a expressão da sequência de ácido nucleico de (a) e opcionalmente

(c) uma sequência de terminação de transcrição.

Preferivelmente, o ácido nucleico que codifica um

polipeptídeo LBD é como definido acima. O termo "sequência de controle" e "sequência de terminação" são como definidos neste.

As Plantas são transformadas com um vetor que compreende qualquer um dos ácidos nucleicos descritos acima. O técnico habilitado estará bem ciente dos elementos genéticos que devem estar presentes no vetor na ordem para transformar, selecionar e propagar de maneira bem sucedida as células hospedeiras contendo a sequência de interesse. A sequência de interesse é operacionalmente ligada a uma ou mais sequências de controle (pelo menos a um promotor).

Vantajosamente, qualquer tipo de promotor, se natural ou sintético, pode ser usado para conduzir a expressão da sequência de ácido nucleico. Um promotor constitutivo é particularmente útil nos métodos. Ver a seção "definições" neste para definições dos vários tipos de promotores.

Deve estar claro que a aplicabilidade da presente invenção não está restrita ao ácido nucleico que codifica o polipeptídeo LBD representado pela SEQ ID N°: 1, nem é a aplicabilidade da invenção restrita à super-expressão de um ácido nucleico que codifica o polipeptídeo LBD quando conduzida por um promotor constitutivo.

O promotor constitutivo é preferivelmente um promotor GOS2, preferivelmente um promotor GOS2 de arroz. Ainda preferivelmente o promotor constitutivo é representado por uma sequência de ácido nucleico substancialmente similar à SEQ ID N°: 10, mais preferivelmente o promotor constitutivo é como representado pela SEQ ID N°: 10. Ver a seção "definições" neste para exemplos adicionais de promotores constitutivos.

Opcionalmente, uma ou mais sequências terminadoras podem ser usadas na construção introduzida em uma planta. Os Elementos reguladores adicionais podem incluir elementos de transcrição bem como de tradução. Aqueles habilitados na técnica estarão cientes de sequências terminadoras e intensificadoras que podem ser adequados para o uso na

realização da invenção. Uma sequência de íntron também pode ser adicionada à região 5' não traduzida (UTR) ou na sequência codificadora para aumentar a quantidade da mensagem madura que acumula-se no citosol, como descrito na seção de definições. Outras sequências de controle (além do promotor, intensificador, silenciador, sequências de íntron, regiões 3'UTR e/ou 5'UTR) podem ser elementos estabilziadores de proteína e/ou RNA. Tais sequências devem ser conhecidas ou podem ser facilmente obtidas por uma pessoa habilitada na técnica.

As construções genéticas da invenção ainda podem incluir uma origem de sequência de replicação que é requerida para a manutenção e/ou replicação em um tipo celular específico. Um exemplo ocorre quando a construção genética é requerida ser mantida em uma célula bacteriana como um elemento genético epissômico (por exemplo, molécula de plasmídeo ou cosmídeo). As origens preferidas de replicação incluem, mas não são limitadas a fl-ori e colEl.

Para a detecção da transferência bem sucedida das sequências de ácido nucleico como usado nos métodos da invenção e/ou seleção de plantas transgênicas que compreendem estes ácidos nucleicos, é vantajoso usar genes marcadores (ou genes repórteres). Portanto, a construção genética pode conter opcionalmente um gene de marcador selecionável. Os marcadores selecionáveis são descritos em mais detalhes na seção "definições" neste.

É conhecido que na integração estável ou transitória de ácidos nucleicos em células de planta, apenas uma minoria das células absorve o DNA estranho e, se desejado, integra-o em seu genoma, dependendo do vetor de expressão usado e da técnica de transfecção usada. Para identificar e selecionar estes integrantes, um gene codificador para um marcador selecionável (tal como os descritos acima) é usualmente introduzido nas células hospedeiras junto com o gene de interesse. Por exemplo,estes marcadores podem ser usados em mutantes em que estes genes não são

funcionais, por exemplo, pela anulação pelos métodos convencionais. Além disso, moléculas de ácido nucleico que codificam um marcador selecionável podem ser introduzidas em uma célula hospedeira no mesmo vetor que compreende a sequência que codifica os polipeptídeos da invenção ou usadas nos métodos da invenção ou ainda em um vetor separado. As células que foram estavelmente transfectadas com o ácido nucleico introduzido podem ser identificadas, por exemplo, pela seleção (por exemplo, células que integraram o marcador selecionável sobrevivem visto que as outras células morrem). Os genes marcadores podem ser removidos ou excisados da célula transgênica uma vez que estes não são mais necessários. As técnicas para a remoção do gene marcador são bem conhecidas na técnica, as técnicas úteis são descritas acima na seção de definições.

A invenção também fornece um método para a produção de plantas transgênicas tendo traços intensificados relacionados com o rendimento com relação às plantas de controle, que compreende introdução e expressão em uma planta de qualquer ácido nucleico que codifica um polipeptídeo LBD como definido neste acima.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo traços aumentados intensificados relacionados com o rendimento, particularmente biomassa aumentada de broto e rendimento aumentado de semente, cujo método compreende:

(i) introduzir e expressar em uma planta ou célula de planta um ácido nucleico que codifica o polipeptídeo LBD e

(ii) cultivar a célula de planta sob condições que promovam crescimento e desenvolvimento de planta.

O ácido nucleico de (i) pode ser qualquer um dos ácidos nucleicos capazes de codificar um polipeptídeo LBD como definido neste.

O ácido nucleico pode ser introduzido diretamente em uma

célula de planta ou na planta por si só (incluindo introdução em um tecido, órgão ou qualquer outra parte de uma planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é preferivelmente introduzido em uma planta por transformação. O termo "transformação" é descrito em mais detalhes na seção "definições" neste.

As células de planta geneticamente modificadas podem ser regeneradas por intermédio de todos os métodos com os quais o trabalhador habilitado está familiarizado. Os métodos adequados podem ser encontrados nas publicações mencionadas acima por S.D. Kung e R. Wu, Potrykus ou Hofgen e Willmitzer.

No geral após a transformação, as as células de planta ou os agrupamentos de célula são selecionados quanto à presença de um ou mais marcadores que são codificados por genes expressíveis em plantas co-transferidas com o gene de interesse, seguindo que o material transformado é regenerado em uma planta total. Para selecionar as plantas transformadas, o material vegetal obtido na transformação é, como uma regra, submetido à condições seletivas de modo que as plantas transformadas possam ser distinguidas das plantas não transformadas. Por exemplo, as sementes obtidas da maneira descrita acima podem ser plantadas e, após um período de desenvolvimento, submetido a uma seleção adequada por pulverização. Uma possibilidade adicional consiste no desenvolvimento de sementes, se apropriado após a esterilização, em placas de ágar usando-se um agente de seleção adequado de modo que apenas as sementes transformadas possam se desenvolver em plantas. Alternativamente, as plantas transformadas são avaliadas quanto à presença de um marcador selecionável tal como os descritos acima.

Seguindo a transferência e a regeneração de DNA, as plantas putativamente transformadas também podem ser avaliadas, por exemplo, usando-se análise de Southern, quanto à presença do gene de interesse,

número de cópias e/ou organização genômica. Alternativa ou adicionalmente, os níveis de expressão do DNA recentemente introduzido pode ser monitorado usando-se análise de Northern e/ou Western, ambas as técnicas sendo bem conhecidas pelas pessoas tendo habilidade comum na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de meios, tais como pela propagação clonal ou técnicas de criação clássicas. Por exemplo, uma planta transformada de primeira geração (ou Ti) podem ser tornadas independentes e transformantes de segunda geração homozigotos (ou T2) selecionados e as plantas T2 aidna podem ser propagadas através de técnicas de criação clássicas. Os organismos gerados podem tomar uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e de células não transformadas; transformantes clonais (por exemplo, todas as células transformadas contém o cassete de expressão); enxertos de tecidos transformados e não transformados (por exemplo, em plantas, um estoque de raiz transformado a um descendente não transformado).

A presente invenção claramente estende-se a qualquer célula de planta ou planta produzida por qualquer um dos métodos descritos neste e a todas as partes de plantas e seus propagules. A presente invenção estende-se ainda para abranger a progênie de uma célula, tecido, órgão ou planta total transformadas ou transfectadas primárias que foram produzidos por qualquer um dos métodos mencionados acima, o único requerimento sendo que a progênie apresente as mesmas características genotípicas e/ou fenotípicas como aqueles produzidos pelo precursor nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucleico isolado que codifica um polipeptídeo LBD como definido neste acima. As células hospedeiras preferidas de acordo com a invenção são células de planta. As plantas hospedeiras para os ácidos nucleicos ou para o

vetor usadas no método de acordo com a invenção, o cassete de expressão ou construção ou vetor são, em princípio, vantajosamente todas as plantas, que são capazes de sintetizar os polipeptídeos usados no método inventivo.

Os métodos da invenção são vantajosamente aplicáveis a qualquer planta. As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular, plantas monocotiledôneas e dicotiledôneas incluindo forragem e legumes de forragem, plantas ornamentais, lavouras de alimentos, árvores ou arbustos. De acordo com a forma de realização preferida da presente invenção, a planta é uma planta de cultivo. Os exemplos de plantas de cultivo incluem soja, girassol, canola, alfafa, semente de colza, algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta monocotiledônea. Os exemplos de plantas monocotiledôneas incluem cana de açúcar. Mais preferivelmente a planta é um cereal. Os Exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

A invenção também estende-se a partes colhíveis de uma planta tais como mas não limitadas a sementes, folhas, frutas, flores, caules, rizomas, tubérculos e bulbos. A invenção, além disso, diz respeito a produtos derivados, preferivelmente, derivados diretamente de uma parte coletável de uma tal planta, tais como grânulos ou pós secos, óleo, gorura e ácidos graxos ou proteínas.

De acordo com uma característica preferida da invenção, a expressão modulada é a expressão aumentada. Os métodos para aumentar a expressão de ácidos nucleicos ou genes ou produtos de gene, são bem documentados na técnica e os exemplos são fornecidos na seção de definições.

Como mencionado acima, um método preferido para modular (preferivelmente, aumentar) a expressão de um ácido nucleico que codifica um polipeptídeo LBD é pela introdução e expressão em uma planta de um

ácido nucleico que codifica um polipeptídeo LBD; entretanto os efeitos de realizar o método, isto é, intensificado os traços relacionados com o rendimento também pode ser atingido usando-se outras técnicas bem conhecidas, que incluem mas não limitado a rotulação de ativação de T-DNA, TILLING, recombinação homóloga. Uma descrição destas técnicas é fornecida na seção de definições.

A presente invenção também abrange o uso de ácidos nucleicos que codificam polipeptídeos LBD como descrito neste e uso destes polipeptídeos LBD na intensificação de qualquer um dos traços relacionados já mencionados com o rendimento em plantas.

Ácidos nucleicos que codificam polipeptídeo LBD descritos neste ou os polipeptídeos LBD por si só, podem encontrar uso nos programas de criação em que um marcador de DNA é identificado podendo estar geneticamente ligado a um gene que codifica o polipeptídeo LBD. Os ácidos nucleicos/genes ou os polipeptídeos LBD por si sós podem ser usados para definir um marcador molecular. Este DNA ou marcador de proteína podem ser então usados nos programas de criação para selecionar plantas tendo traços intensificados relacionados com o rendimento como definido neste acima nos métodos da invenção.

As variantes alélicas de um ácido nucleico que codifica o polipeptídeo LBD/gene também podem encontrar uso em programas de criação auxiliados por marcador. Tais programas de criação algumas vezes requerem a introdução de variação alélica pelo tratamento mutagênico das plantas, usando-se por exemplo mutagênese por EMS; alternativamente, o programa pode iniciar com uma coleção de variantes alélicas da denominada origem "natural" não intencionalmente. A identificação de variantes alélicas então acontece, por exemplo, por PCR. Isto é seguido por uma etapa para a seleção de variantes alélicas superiores da sequência em questão e que dão rendimento aumentado. A seleção é tipicamente realizada pelo

monitoramento do desempenho de desenvolvimento de plantas contendo variantes alélicas da sequência em questão. O desempenho de desenvolvimento pode ser monitorado em uma estufa ou no campo. As etapas opcionais adicionais incluem o cruzamento de plantas em que a variante alélica foi identificada com uma planta. Isto pode ser usado, por exemplo, para fabricar uma combinação de características fenotípicas interessantes.

Os ácidos nucleicos que codificam os polipeptídeos LBD também podem ser usados como sondas para mapear genética e fisicamente os genes que estes são uma parte de e como marcadores para traços ligados àqueles genes. Tal informação pode ser útil na criação de plantas a fim de desenvovler linhas com fenotipos desejados. Tal uso de ácido nucleico que codifica os polipeptídeos LBD requer apenas uma sequência de ácido nucleico de pelo menos 15 nucleotídeos de comprimento. O ácido nucleico que codifica os polipeptídeos LBD pode ser usado como marcadores de polimorfismo de comprimento de fragmento (RFLP). Southern blots (Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico de planta digerida por restrição pode ser sondada com os ácidos nucleicos que codificam LBD. Os padrões de formação de banda resultantes podem ser então submetidos a análises genéticas usando-se programas de computador, tais como MapMaker (Lander et al. (1987) Genomics 1: 174-181) a fim de construir um mampa genético. Além disso, os ácidos nucleicos podem ser usados para Southern blots de sonda contendo DNAs genômicos tratados com endonuclease de restrição de uma série de indivíduos que representam precursores e progênie de um cruzamento genético definido. A segregação dos polimorfismos de DNA é observado e usado para calcular a posição do ácido nucleico que codifica o polipeptídeo LBD no mapa genético previamente obtido usando-se esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

A produção e o uso de sondas derivadas de gene vegetal para o

uso no mapeamento genético são descritos em Bernatzky and Tanksley (1986) Plant Mol. Biol. Reporter 4: 37-41. As numerosas publicações descrevem o mapeamento genético de clones de cDNA específicos usando-se a metodologia resumida acima ou variações desta. Por exemplo, populações intercruzamento F2, populações de retrocruzamento, populações aleatoriamente comparadas, linhas isogênicas próximas e outras séries de indivíduos podem ser usado para o mapeamento. Tais metodologias são bem conhecidas por aqueles habilitados na técnica.

As sondas de ácido nucleico também podem ser usadas para o mapeamento físico (isto é, colocação de sequências em mapas físicos; ver Hoheisel et al. em: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346 e referências citadas neste).

Em uma outra forma de realização, as sondas de ácido nucleico pode ser usadas no mapeamento de hibridização por fluorescência direta in situ (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora os métodos correntes de mapeamento FISH favoreçam o uso de clones grandes (diversos kb a diversas centenas de kb; see Laan et al. (1995) Genome Res. 5:13-20), melhoras na sensibilidade podem permitir o desempenho do mapeamento FISH usando-se sondas mais curtas.

Uma variedade de métodos com base na amplificação de ácido nucleico para o mapeamento genético e físico pode ser realizada usando-se o ácido nucleicos. Os exemplos incluem a amplificação específica de alelo (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332), ligação específica de alelo (Landegren et al. (1988) Science 241:1077-1080), reações de extensão de nucleotídeo (Sokolov (1990) Nucleic Acid Res. 18:3671), Mapeamento Híbrido por Radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e Mapeamento Happy (Dear and Cook (1989) Nucleic Acid Res. 17:6795-6807). Para estes métodos, a sequência de um ácido nucleico é

usada para projetar e produzir pares de iniciador para o uso na reação de amplificação ou nas reações de extensão de iniciador. O projeto de tais iniciadores é bem conhecido por aqueles habilitados na técnica. Em métodos que utilizam mapeamento genético com base em PCR, pode ser necesário identificar diferenças de sequência de DNA entre os precursores do mapeamento do cruzamento na região que corresponde à presente sequência de ácido nucleico. isto, entretanto, não é necessário para os métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo traços intensificados relacionados com o rendimento, como descrito anteriormente. Estes traços também podem ser combinados com outros traços vantajosamente econômicos, tais como traços de intensificação de rendimento adicionais, tolerância a outras tensões bióticas ou abióticas, traços que modificam várias características arquitetônicas e/ou características bioquímicas e/ou fisiológicas.

II. polipeptídeo JMJC (JUMONJI-C)

Surpreendentemente, foi observado agora que modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo JMJC dá plantas tendo traços intensificados relacionados com o rendimento com relação às plantas de controle. De acordo com uma primeira forma de realização, a presente invenção fornece um método para intensificar traços relacionados com o rendimento em plantas com relação às plantas de controle, que compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo JMJC.

Um método preferido para modular (preferivelmente, aumentar) expressão de um ácido nucleico que codifica um polipeptídeo JMJC é pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo JMJC.

Qualquer referência a seguir a uma "proteína útil nos métodos

da invenção" é determinada entender um polipeptídeo JMJC como definido neste. Qualquer referência a seguir a um "ácido nucleico útil nos métodos da invenção" é determinada entender um ácido nucleico capaz de codificar um tal polipeptídeo JMJC. O ácido nucleico a ser introduzido em uma planta (e, portanto, útil na realização dos métodos da invenção) é qualquer ácido nucleico que codifica o tipo de proteína, que será descrito agora, a seguir também denominado "JMJC ácido nucleico" ou "gene JMJC".

Um polipeptídeo JMJC como definido neste refere-se a um polipeptídeo que compreende pelo menos um domínio JmjC.

O domínio JmjC é uma sequência conservada encontrada em proteínas de organismo procarióticos e eucarióticos. Os domínios JmjC têm em média 111 aminoácidos de comprimento. Tipicamente, o comprimento de um domínio JmjC pode variar entre 25 e 200 aminoácidos embora versões mais curtas e mais longas possam ser possíveis. Os domínios JmjC são preditos serem metaloenzimas que adotam a dobra de cupina e são candidatos para enzimas que regulam a remodelagem de cromatina (Clissold et al. Trends Biochem Sci. 2001 Jan;26(1):7-9).

Os polipeptídeos JMJC podem ser observados em bancos de dados especialziados, tais como Pfam, (Finn et al. Nucleic Acids Research (2006) Database Issue 34:D247-D251). O Pfam compila uma grande coleção de alinhamentos de sequência múltiplos e modelos de Markov escondidos (HMM) que cobrem os domínios e famílias de muito comuns e está disponível através do Sanger Institute no Reino Unido.

O limiar de corte de união do domínio JmjC no Pfam modelo HMM_fs é de 16,0 e de -8,0 no modelo HMM_Is. Comparações de confiança como consideradas no banco de dados Pfam são aquelas de registro mais alto do que o limiar de corte de união. Entretanto, comparações potenciais, que compreendem domínios JmjC verdadeiros, ainda podem estar sob o corte de união. Preferivelmente um polipeptídeo JMJC é uma proteína tendo um o

mais domínios em sua sequência que excede o corte de união da família de domínio de proteína Pfam PF02373, jumonji, jmjC.

Alternativamente, um domínio JmjC em um polipeptídeo pode ser identificado pela realização de uma comparação de sequência com polipeptídeos conhecidos que compreendem um domínio JmjC e que estabelece a similaridade na região do domínio JmjC. As sequências podem ser alinhadas usando-se qualquer um dos métodos bem conhecidos na técnica, tais como algoritmos Blast e a probabilidade para o alinhamento ocorrer com uma dada sequência é tomada como base para a identificação de polipeptídeos similares. Um parâmetro que é tipicamente usado para a representação de tal probabildade nas comparações em pares é denominado valor e. O valor e descreve como frequentemente, é esperado que um dado registro ocorra de maneira aleatória. O corte do valor e pode ser tão alto quanto 1,0. Tipicamente os alinhamentos com probabilidade alta de ocorrer têm um valor e menor do que 0,1, 0,01, 0,001, 1.e-05, 1.e-10, 1.e-15, 1.e-20, 1.e-25, 1.e-50, 1.e-75, 1.e-100 e 1.e-200. Preferivelmente, os polipeptídeos JMJC da invenção compreendem uma sequência tendo na ordem crescente de preferência um valor e menor do que 0,1, 0,01, 0,001, 1.e-05, 1.e-10, 1.e-15, 1.e-20, 1.e-25, 1.e-50, 1.e-75, 1.e-100 e 1.e-200 para um alinhamento com o domínio JmjC observado em um polipeptídeo JMJC conhecido.

Os exemplos de polipeptídeos JMJC são dados na Tabela B1. As coordenadas de aminoácido para os domínios JmjCs como presente nos polipeptídeos JMJC representativos de origem de dicotiledônea são dadas na Tabela B4. Um exemplo de domínio JmjC é dado dada na SEQ ID N°: 78.

Uma identidade de sequência entre os domínios domínio JmjC é tipicamente bixa, em média 23 % e pode ser tão baixo quanto tão baixo quanto 10 %. Os polipeptídeos JMJC preferidos da invenção são aqueles tendo na ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, ou mais identidade de sequência a SEQ

ID N°: 78 ou a qualquer um dos domínios JmjC compreendidos nos polipeptídeos JMJC representado pela SEQ ID N°: 84; SEQ ID N°: 86; SEQ ID N°: 96; SEQ ID N°: 98; SEQ ID N°: 104; SEQ ID N°: 108; SEQ ID N°: 110; SEQ ID N°: 112; SEQ ID N°: 114; SEQ ID N°: 116; SEQ ID N°: 118; SEQ ID N°: 120; SEQ ID N°: 122; SEQ ID N°: 124; SEQ ID N°: 128; SEQ ID N°: 130; SEQ ID N°: 132; e SEQ ID N°: 134, cujas coordenadas de aminoácido são dadas na Tabela B4.

Além do domínio JmjC, os polipeptídeos JMJC podem compreender opcionalmente outros motivos de sequência altamente conservados, tais como aqueles representados SEQ ID N°: 79, SEQ ID N°: 80, SEQ ID N°: 81, que são encontrados na SEQ ID N°: 74. Além disso, os polipeptídeos da invenção podem compreender uma sequência conservada encontradas em oxigenases que estavam envolvidas na coordenação de cátions de ferro, aqui representado por HXD(V) ou EXnH (SEQ ID N°: 82), em que "X" representa qualquer aminoácido e "n" representa o número de X-resíduos que variam entre 1 e 5, ambos incluídos.

Um polipeptídeo JMJC preferido da invenção refere-se a qualquer polipeptídeo que compreenda um domínio JmjC e opcionalmente tendo na ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % ou mais identidade de sequência a qualquer um ou mais dos motivos de sequência conservados como representado pela SEQ ID N°: 79, SEQ ID N°: 80, SEQ ID N°: 81 e SEQ ID N°: 82.

Evidência de domínio trocando na família JUMONJI de proteínas foi relatado, (Balciunas and Ronne. Trends Biochem Sci 2000;25:274-276). Consequentemente, além do domínio JmjC, os polipeptídeos JMJC tipicamente contém um ou mais diferentes tipos de outros domínios conservados conhecidos. Relevantes aos polipeptídeos da invenção são os polipeptídeos JMJC que além do domínio JmjC compreende outros

domínios conservados, que estão presumivelmente envolvidos na ligação de DNA e atividades de transcrição, tais como dedos de zinco e/ou na interação de proteína e revoluçõesde proteína, tais como F-box e domínios tipo dedo-RING de zinco.

Portanto o polipeptídeo JMJC útil nos métodos da invenção compreende um domínio JmjC e opcionalmente um ou mais dos seguintes domínios conservados: JmJN (número de acessão pfam PF02375); dedo de zinco C5HC2 (número de acessão pfam: PF02928), domínio rico em FY, terminal N (número de acessão InterPro: IPR003888) e domínio rico em FY, terminal C (número de acessão InterPro: IPR003889), um dedo de zinco tipo FYVE/PHD-Zn (Dedo de zinco InterPro: IPR011011), um dedo de zico tipo C2H2 (número de acessão pfam: PF00096), um dedo de zinco – tipo RING (zf-C3HC4) (número de acessão pfam PF00097) e um domínio F-box (número de acessão pfam: PF00646).

Preferivelmente o polipeptídeo JMJC da invenção compreende uma sequência tendo na ordem crescente de preferência de 50 %, 60 %, 70 %, 75 %, 60 %, 85 %, 90 %, 92 %, 94 %, 96 %, 98 % ou mais identidade de sequência a SEQ ID N°: 74, SEQ ID N°: 86, SEQ ID N°: 94, SEQ ID N°: 104, SEQ ID N°: 122, SEQ ID N°: 134, SEQ ID N°: 136, SEQ ID N°: 140, SEQ ID N°: 142 e SEQ ID N°: 148.

Preferivelmente, a sequência de polipeptídeo que quando usada na construção de uma árvore filogenética, tal como o descrito na Figura 8, agrupamentos com o grupo de polipeptídeos JMJC (Takeuchi et al. Dev Dyn. 2006 235(9): 2449-59) que compreende a sequência de aminoácido representada pela SEQ ID N°: 74 em vez de qualquer outro grupo.

O termo "domínio" e "motivo" é definido na seção "definições" neste. Bancos de dados especialistas existem para a identificação de domínios, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-

244, InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318, Prosite (Bucher and Bairoch (1994), Uma sintaxe de perfil generalizada para motivos de sequências biomoleculares e sua função em interpretação de sequência automática. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004) ou Pfam (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002). Uma série de ferramentas para a análise *in silico* de sequências de proteínas está disponível no servidor ExPASy proteomics (Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31:3784-3788(2003)). Os domínios também podem ser identificados usando-se técnicas de rotina, tal como pelo alinhamento de sequência.

Os métodos para o alinhamento de sequências para comparação são bem conhecidos na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA and TFASTA. GAP usa o algoritmo de Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar o alinhamento global (isto é medindo-se as sequências completas) de duas sequências que maximizam o número de equiparações e minimiza o número de fendas. o algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-10) calcula a identidade de sequência percentual e realiza uma análise estatística da similaridade entre as duas sequências. O software para realizar análise BLAST está publicamente disponível através do National Centre for Biotechnology Information (NCBI). Os homólogos podem ser facilmente identificados usando-se, por exemplo, o algoritmo de alinhamento de sequência múltipla ClustalW (versão 1.83), com os parâmetros de alinhamento default e um método de registro em porcentagem. As porcentagens globais de similaridade e de identidade também podem ser

determinadas usando-se um dos métodos disponíveis no pacote de software MatGAT (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando-se proteína ou sequências de DNA.). A edição manual menor pode ser realizada para otimizar o alinhamento entre os motivos conservados, como deve estar evidente para uma pessoa habilitada na técnica. Além disso, em vez de usar sequências de comprimento total para a identificação de homólogos, domínios específicos também podem ser usados. Os valores de identidade de sequência podem ser determinadas no aminoácido inteiro ou sequência de aminoácido ou em domínios selecionados ou motivos conservados, usando-se os programas mencionados acima usando-se os parâmetros default.

Além disso, os polipeptídeos JMJC podem ter tipicamente atividade de hidroxilases de proteína, em particular, atividade de peptídeo-aspartate beta-dioxigenase (PABD) (EC 1.14.11.16). a reação catalisada é: peptídeo L-aspartato + 2-oxoglutarato + 0(2) <=> peptídeo 3-hidróxi-L-aspartato + succinato + CO(2). Os cofatores na reação são ferro e 2-Oxogluatarato. As ferramentas e técnicas para medir a atividade de PABD são bem conhecidas na técnica (ver, por exemplo, Lavaissiere et al. J Clin Invest. 1996;98(6):1313-23; Linke et al. J Biol Chem. 2004;279(14):14391-7; Lee, et al. J. Biol. Chem. 278:7558-7563; 2003; Lando et al. Genes Dev. 16:1466-1471; 2002). As oxigenases dependentes de ferro (II)/2-oxoglutarato (2-0G) catalsiam as reações oxidativas em várias reações metabólicas. Recentemente foi proposto que os polipeptídeos JMJC também possam ter atividade de histona desmetilase (Trewick et al. 2005).

A presente invenção é ilustrada transformando-se as plantas com a sequência de ácido nucleico representada pela SEQ ID N°: 73, que codifica a sequência de polipeptídeo da SEQ ID N°: 74. Entretanto, o desempenho da invenção não está restrito a estas sequências; os métodos da invenção podem ser vantajosamente realizados usando-se qualquer ácido

nucleico que codifica JMJC ou polipeptídeo JMJC como definido neste.

Os exemplos de ácidos nucleicos que codificam polipeptídeos JMJC são dadas na Tabela B1 do Exemplo 12 neste. Tais ácidos nucleicos são úteis na realização dos métodos da invenção. As sequências de aminoácido dadas na Tabela B1 do Exemplo 12 são sequências de exemplos de ortólogos e parálogos do polipeptídeo JMJC representado pela SEQ ID N°: 74, os termos "ortólogos" e "parálogos" sendo como definido neste. Ortólogos e parálogos adicionais podem ser facilmente definidos pela realização de uma denominada pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envovle submeter uma sequência em dúvida ao BLAST (por exemplo usando-se qualquer uma das sequências listadas na Tabela B1 do Exemplo 12) contra qualquer banco de dados de sequência, tal como os bancos de dados NCBI publicamente disponíveis. BLASTN ou TBLASTX (usando-se os valores default padrão) são, no geral, usados quando começa-se a partir de uma sequência de nucleotídeo e BLASTP ou TBLASTN (usando-se os valores default padrão) quando começa-se a partir de uma sequência de proteína. Os resultados de BLAST podem ser opcionalmente filtrados. As sequências de comprimento total de cada um dos resultados filtrados ou resultados não filtrados são então submetidos ao BLAST novamente (segundo BLAST) contra sequências do organismo a partir do qual a sequência em questão é derivada (quando a sequência em questão é SEQ ID N°: 73 ou SEQ ID N°: 74, o segundo BLAST portanto, deve ser contra sequências de Arabidopsis). Os resultados do primeiro e segundo BLASTs são então comparados. Um parálogo é identificado se um acerto de classificação alta do primeiro blast for da mesma espécie como a partir da qual a sequência em questão é derivada, submetendo-se ao BLAST novamente depois idealmente resulta na sequência em questão entre os acertos mais altos; um ortólogo é identificado se um acerto de classificação alta no primeiro BLAST não for da mesma espécie como a partir da qual a sequência em questão é derivada e

preferivelmente resulta no BLAST novamente na sequência em questão estando entre os acertos mais altos.

Os acertos de classificação alta são aqueles tendo um valor E baixo. Quanto mais baixo o valor E, mais significante o registro (ou, em outras palavras, menor a chance que o acerto seja encontrado por chance).A computação do valor E é bem conhecida na técnica. Além dos valores E, comparações são registradas pela identidade de porcentagem. Identidade de porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácidos nucleicos comparadas (ou polipeptídeo) em um comprimento particular. No caso de famílias grandes, ClustalW pode ser usado, seguido por uma árvore de união próxima, para ajudar a visualizar o agrupamento de genes relacionados e para identificar ortólogos e parálogos.

As variantes de ácido nucleico também podem ser úteis na prática dos métodos da invenção. Os exemplos de tais variantes incluem ácidos nucleicos que codificam homólogos e derivados de qualquer uma das sequências de aminoácido dadas na Tabela B1 do Exemplo 12, os termos "homólogo" e "derivado" sendo como definido neste. Também são úteis nos métodos da invenção os ácidos nucleicos que codificam homólogos e derivados de ortólogos ou parálogos de qualquer uma das sequências de aminoácido dadas na Tabela B1 do Exemplo 12. Homólogos e derivados úteis nos métodos da presente invenção têm substancialmente a mesma atividade biológica e funcional como a proteína não modificada da qual estes são derivados.

As variantes de ácido nucleico adicionais úteis na prática dos métodos da invenção incluem porções de ácidos nucleicos que codificam polipeptídeos JMJC, ácidos nucleicos que hibridizam-se a ácidos nucleicos que codificam polipeptídeos JMJC, variantes de união de ácidos nucleicos que codificam polipeptídeos JMJC, variantes alélicas de ácidos nucleicos que codificam polipeptídeos JMJC e variantes de ácidos nucleicos que codificam

polipeptídeos JMJC obtidos pelo embaralhamento de gene. Os termos sequência de hibridização, variante de união, variante alélica e embaralhamento de gene são como descritos neste.

Os ácidos nucleicos que codificam polipeptídeos JMJC não necessitam ser ácidos nucleicos de comprimento total, visto que o desempenho dos métodos da invenção não contam com o uso de sequências de comprimento total de ácido nucleico. De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta uma porção de qualquer uma das sequências de ácido nucleico dadas na Tabela B1 do Exemplo 12 ou uma porção de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de aminoácido dadas na Tabela B1 do Exemplo 12.

Uma porção de um ácido nucleico pode ser preparada, por exemplo, realizando-se uma ou mais anulações no ácido nucleico. As porções podem ser usadas na forma isolada ou estas podem ser fundidas a outras sequências codificadoras (ou não codificadoras) a fim de, por exemplo, produzir uma proteína que combine diversas atividades. Quando fundido a outras sequências codificadoras, o polipeptídeo resultante produzido na tradução pode ser maior do que aquele predito para a porção de proteína.

As porções úteis métodos da invenção, codificam um polipeptídeo JMJC como definido neste têm substancialmente a mesma atividade biológica como as sequências de aminoácido dadas na Tabela B1 do Exemplo 12. Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos dados na Tabela B1 do Exemplo 12 ou é uma porção de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das sequências de aminoácido dadas na Tabela B1 do Exemplo 12. Preferivelmente a porção is pelo menos 100, 200, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500,

5000 nucleotídeos de comprimento consecutivo, os nucleotídeos consecutivos sendo de qualquer uma das sequências de ácido nucleico dadas na Tabela B1 do Exemplo 12 ou de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das sequências de aminoácido dadas na Tabela B1 do Exemplo 12. Mais preferivelmente a porção é uma porção do ácido nucleico da SEQ ID N°: 73. Preferivelmente, a porção codifica uma sequência de aminoácido que quando usada na construção de uma árvore filogenética, tal como o descrito na Figura 8, agrupamentos com o grupo de polipeptídeos JMJC (Takeuchi et al. Dev Dyn. 2006 235(9):2449-59) que compreende a sequência de aminoácido representada pela SEQ ID N°: 74 em vez de qualquer outro grupo.

Uma outra variante de ácido nucleico útil nos métodos da invenção é um ácido nucleico capaz de hibridizar, sob condições de estringência reduzida, preferivelmente sob condições estringentes, com um ácido nucleico que codifica um polipeptídeo JMJC como definido neste ou com uma porção como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta, um ácido nucleico capaz de hibridizar qualquer um dos ácidos nucleicos dados na Tabela B1 do Exemplo 12 ou que compreende introduzir e expressar em uma planta, um ácido nucleico capaz de hibridizar a um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de ácido nucleico dadas na Tabela B1 do Exemplo 12.

As sequências de hibridização úteis nos métodos da invenção codificam um polipeptídeo JMJC como definido neste têm substancialmente a mesma atividade biológica como as sequências de aminoácido dadas na Tabela B1 do Exemplo 12. Preferivelmente, a sequência de hibridização é capaz de hibridizar a qualquer um dos ácidos nucleicos dados na Tabela B1

do Exemplo 12 ou a uma porção de qualquer destas sequências, uma porção sendo como definida acima; or em que a sequência de hibridização é capaz de hibridizar a um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das sequências de aminoácido dadas na Tabela B1 do Exemplo 12. Mais preferivelmente, a sequência de hibridização é capaz de hibridizar a um ácido nucleico como representado pela SEQ ID N°: 73 ou a uma porção deste.

Preferivelmente, a sequência de hibridização codifica uma sequência de aminoácido que quando usada na construção de uma árvore filogenética, tal como o descrito na Figura 8, agrupamentos com o grupo de polipeptídeos JMJC (Takeuchi et al. Dev Dyn. 2006 235(9):2449-59) que compreende a sequência de aminoácido representada pela SEQ ID N°: 74 em vez de qualquer outro grupo.

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante de união que codifica um polipeptídeo JMJC como definido neste acima, a variante de união sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta a variante de união de qualquer uma das sequências de ácido nucleico dadas na Tabela B1 do Exemplo 12 ou a variante de união de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de aminoácido dadas na Tabela B1 do Exemplo 12. um exemplo de uma variante unida do gene que codifica SEQ ID N°: 74 é representado pela SEQ ID N°: 73 e SEQ ID N°: 83 (que codifica a SEQ ID N°: 84).

As variantes de união preferidas são variantes de união de um ácido nucleico representado pela SEQ ID N°: 73 ou a variante de união de um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 74. Preferivelmente, a sequência de aminoácido codificada pela variante de união,

quando usado na construção de uma árvore filogenética, tal como o descrito na Figura 8, agrupamentos com o grupo de polipeptídeos JMJC (Takeuchi et at. Dev Dyn. 2006 235(9):2449-59) que compreende a sequência de aminoácido representada pela SEQ ID N°: 74 qualquer outro grupo.

Uma outra variante de ácido nucleico útil na realização dos métodos da invenção é uma variante alélica de um ácido nucleico que codifica um polipeptídeo JMJC como definido neste acima, uma variante alélica sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta uma variante alélica de qualquer um dos ácidos nucleicos dados na Tabela B1 do Exemplo 12 ou que compreende introduzir e expressar em uma planta uma variante alélica de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de aminoácido dadas na Tabela B1 do Exemplo 12.

As variantes alélicas úteis nos métodos da presente invenção têm substancialmente a mesma atividade biológica como o polipeptídeo JMJC da SEQ ID N°: 74 qualquer um dos aminoácidos descritos na Tabela B1 do Exemplo 12. As Variante alélicas existem em estado natural e abrangidas dentro dos métodos da presente invenção são o uso destes alelos naturais. Preferivelmente, a variante alélica é uma variante alélica da SEQ ID N°: 73 ou uma variante alélica de um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 74. Preferivelmente, a sequência de aminoácido codificada pela variante alélica, quando usada na construção de uma árvore filogenética, tal como o descrito na Figura 8, agrupa-se com os polipeptídeos JMJC (Takeuchi et al. Dev Dyn. 2006 235(9):2449-59) que compreende a sequência de aminoácido representada pela SEQ ID N°: 74 em vez de qualquer outro grupo.

A mistura de gene e a evolução direcionada podem ser usadas

para gerar variantes de ácidos nucleicos que codificam polipeptídeos JMJC como definido acima; o termo "mistura de gene" sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta uma variante de qualquer uma das sequências de ácido nucleico dadas na Tabela B1 do Exemplo 12 ou que compreende introduzir e expressar em uma planta uma variante de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de aminoácido dadas na Tabela B1 do Exemplo 12, cujo ácido nucleico variante é obtido pelo embaralhamento de gene.

Vantajosamente, a presente invenção fornece até agora sequências de ácido nucleico e de polipeptídeo JMJ desconhecidas.

De acordo com uma outra forma de realização da presente invenção, é fornecida uma molécula de ácido nucleico isolada que compreende:

(i) um ácido nucleico representado pela SEQ ID N°: 169;

(ii) um ácido nucleico ou fragmento do mesmo que é complementar à SEQ ID N°: 169;

(iii) um ácido nucleico que codifica um polipeptídeo JMJC tendo, na ordem crescente de preferência, pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência à sequência de aminoácido dada na SEQ ID N°: 170;

(iv) um ácido nucleico capaz de hibridizar sob condições estringentes a qualquer um dos ácidos nucleicos dados em (i), (ii) ou (iii) acima.

De acordo com uma outra forma de realização da presente invenção, portanto, é fornecido um polipeptídeo isolado que compreende:

(i) uma sequência de aminoácido tendo, na ordem crescente de preferência, pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %,

98 %, 99 % ou 100 % ou mais identidade de sequência à sequência de aminoácido dada na SEQ ID N°: 170;

(ii) derivados de qualquer uma das sequências de aminoácido dadas em (i).

Preferivelmente, a sequência de aminoácido codificada pelo ácido nucleico variante obtidos pelo embaralhamento de gene, quando usado na construção de uma árvore filogenética tal como a descrita na Figura 8, agrupamentos com o grupo de polipeptídeos JMJC (Takeuchi et al. Dev Dyn. 2006 235(9):2449-59) que compreende a sequência de aminoácido representada pela SEQ ID N°: 74 em vez de qualquer outro grupo.

Além disso, as variantes de ácido nucleico também podem ser obtidas por mutagênese direcionada ao local. Diversos métodos estão disponíveis para atingir mutagênese direcionada ao local, o mais comum sendo os métodos com base em PCR (Current Protocols in Molecular Biology. Wiley Eds.).

Os ácidos nucleicos que codificam polipeptídeos JMJC podem ser derivados de qualquer fonte natural ou artificial. O ácido nucleico pode ser modificado a partir de sua forma natural na composição e/ou ambiente genômico através da manipulação humana deliberada. Preferivelmente, o ácido nucleico que codifica o polipeptídeo JMJC é de uma planta, ainda preferivelmente de uma planta dicotiledônea, mais preferivelmente da família Brassicaceae, mais preferivelmente o ácido nucleico é de *Arabidopsis thaliana.*

O desempenho dos métodos da invenção dá plantas tendo traços intensificados relacionados com o rendimento. Em particular, desempenho dos métodos da invenção dá plantas tendo rendimento aumentado, especialmente rendimento aumentado de semente com relação às plantas de controle. Os termos "rendimento" e "rendimento de semente" são descritos em mais detalhes na seção "definições" neste.

Referência neste a traços intensificados relacionados com o rendimento é determinada entender um aumento na biomassa (peso) de uma ou mais partes de uma planta, que pode incluir partes (coletáveis) acima do solo e/ou partes (coletáveis) abaixo do solo. Em particular, tais partes colhíveis são sementes e o desempenho dos métodos da invenção resulta em plantas tendo rendimento aumentado com relação ao rendimento de plantas de controle.

Tomando o milho como um exemplo, um aumento no rendimento pode ser manifestado as um ou mais do seguinte: aumento no número de plantas estabelecido por hectare ou acre, um aumento no número de espigas por planta, um aumento no número de fileiras, número de sementes por fileira, peso de semente, peso de semente milhar, comprimento/diâmetro de espiga, aumento na razão de enchimento de semente (que é o número de sementes enchidas dividido pelo número total de sementes e multiplicado por 100), entre outros. Tomando-se o arroz como um exemplo, um aumento no rendimento pode manifestar-se por si só como um aumento em um ou mais do seguinte: número de plantas por hectare ou acre, número de panículas por planta, número de espigas por panícula, número de flores (flósculos) por panícula (que é expressado como uma razão do número de sementes enchidas no número de panícula primárias), aumento na razão de enchimento de semente (que é o número de sementes enchidas dividido pelo número total de sementes e multiplicado por 100), aumento em peso de semente milhar, entre outros.

A presente invenção fornece um método para aumentar o rendimento, especialmente rendimento de semente de plantas, com relação às plantas de controle, cujo método compreende a expressão modular, preferivelmente aumentar a expressão, em uma planta de um ácido nucleico que codifica um polipeptídeo JMJC como definido neste.

Visto que as plantas transgênicas de acordo com a presente

invenção têm rendimento aumentado, é provável que estas plantas apresentem uma taxa de desenvolvimento aumentada (durante pelo menos parte do seu ciclo de vida), com relação à taxa de desenvolvimento de plantas de controle em um estágio correspondente em seu ciclo de vida.

A taxa de desenvolvimento aumentada pode ser específica para uma ou mais partes de uma planta (incluindo sementes) ou pode ser substancialmente por toda a planta. As plantas tendo uma taxa de desenvolvimento aumentada podem ter um ciclo de vida mais curto. O ciclo de vida de uma planta pode ser tomado significar o tempo necessário para o desenvolvimento de uma semente madura seca até o estágio onde a planta produziu sementes maduras secas, similar ao material de partida. Este ciclo de vida pode ser influenciado por fatores tais como vigor precoce, taxa de desenvolvimento, índice de verdor, tempo de florescimento e velocidade de maturação de semente. O aumento em taxa de desenvolvimento pode acontecer em um ou mais estágios no ciclo de vida de uma planta ou substancialmente durante o ciclo de vida de planta total. A taxa de desenvolvimento aumentada durante os estágios precoces no ciclo de vida de uma planta pode refletir o vigor intensificado. O aumento na taxa de desenvolvimento pode alterar o ciclo de coleta de uma planta permitindo que as plantas sejam semeadas depois e/ou coletadas mais cedo do que seria de outra maneira possível (um efeito similar pode ser obtido com tempo de florescimento precoce). Se a taxa de desenvolvimento for suficientemente aumentada, isto pode permitir a semeadura adicional de sementes das mesmas espécies vegetais (por exemplo semeadura e coleta de plantas de arroz seguido pela semeadura e coleta de outras plantas de arroz tudo dentro de um período de desenvolvimento convencional). Similarmente, se a taxa de desenvolvimento for suficientemente aumentada, isto pode permitir a semeadura adicional de sementes de espécies de plantas diferentes (por exemplo a semeadura e a coleta de plantas de milho seguido por, por

exemplo, a semeadura e a coleta opcional de soja, batata ou qualquer outro vegetal adequado). Os períodos de coleta adicionais a partir do mesmo estoque de raiz no caso de algumas plantas de cultivo também pode ser possível. Alteração do ciclo de coleta de uma planta pode levar a um aumento na produção de biomassa anual por acre (devido a um aumento no número de vezes (diz-se em um ano) que qualquer planta particular pode ser desenvolvida e coletada). Um aumento na taxa de desenvolvimento também pode permitir o cultivo de plantas transgênicas em uma área geográfica mais ampla do que suas contrapartes do tipo selvagem, visto que as limitações territoriais para o desenvolvimento de uma lavoura são frequentemente determinados pelas condições ambientais adversas no período de plantio (estação precoce) ou no período de coleta (estação posterior). Tais condições adversas podem ser evitadas se o ciclo de coleta for encurtado. A taxa de desenvolvimento pode ser determinada derivando-se vários parâmetros de curvas de desenvolvimento, tais parâmetros podem ser: T-Mid (o tempo levado pelas plantas para atingir 50 % de seu tamanho máximo) and T-90 (tempo levado pelas plantas para atingir 90 % de seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção dá plantas tendo uma taxa de desenvolvimento aumentada com relação às plantas de controle. Portanto, de acordo com a presente invenção, é fornecido um método for aumentar a taxa de desenvolvimento de plantas, cujo método compreende a expressão modular, preferivelmente aumentar a expressão, em uma planta de um ácido nucleico que codifica um polipeptídeo JMJC como definido neste.

Um aumento no rendimento e/ou taxa de desenvolvimento ocorre se a planta estiver sob condições sem estresse ou se a planta for exposta a várias tensões em comparação com as plantas de controle. Plantas tipicamente respondem à exposição à tensão desenvolvendo-se mais

lentamente. Em condições de tensão grave, a planta pode ainda parar de se desenvolver completamente. Tensão branda por outro lado é definida neste como sendo qualquer tensão à qual uma planta é exposta que não resulta na planta parando de se desenvolver completamente sem sem a capacidade de retomar o crescimento. Tensão branda no sentido da invenção da invenção leva a uma redução no desenvolvimento das plantas submetidas à tensão de menos que 40 %, 35 % ou 30 %, preferivelmente menor do que 25 %, 20 % ou 15 %, mais preferivelmente menor do que 14 %, 13 %, 12 %, 11 % ou 10 % ou menor em comparação com a planta de controle sob condições sem estresse. Devido aos avanços nas práticas agrícolas (irrigação, fertilização e tratamentos com pesticidas) tensões graves não são frequentemente encontradas em plantas de cultivo cultivadas. Como uma consequência, o desenvolvimento comprometido induzido pela tensão branda é frequentemente uma característica indesejada para a agricultura. Tensões brandas são as tensões (ambientais) bióticas e/ou abióticas diárias às quais a planta é exposta. Tensões abióticas podem ocorrer devido à secura ou excesso de água, tensão anaeróbica, tensão ao sal, toxicidade química, tensão oxidativa e calor, frio ou temperaturas congelantes. A tensão abiótica pode ser uma tensão osmótica causada por uma tensão a água (particularmente devido à secura), tensão ao sal, tensão oxidativa ou uma tensão iônica. Tensões bióticas são tipicamente aquelas tensões causadas por patogênio, tal como bactérias, vírus, fungos e insetos.

Em particular, os métodos da presente invenção podem ser realizados sob condições sem estresse ou sob condições de secura branda para dar plantas tendo rendimento aumentado com relação às plantas de controle. Como relatado em Wang et al. (Plant (2003) 218: 1-14), tensão abiótica leva a uma série de mudanças morfológicas, fisiológica, bioquímicas e moleculares que afetam adversamente o desenvolvimento e produtividade da planta. Secura, salinidade, temperaturas extremas e tensão oxidativa são conhecidas

serem interconectadas e podem induzir o desenvolvimento e o dano celular através de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve um grau particularmente alto de "cruzamento" tensão por secura e tensão por salinidade alta. Por exemplo, seca e/ou desalinização são manifestadas primariamente como tensão osmótica, resultando na interrupção da homeostase e a distribuição de íon na célula. Tensão oxidativa, que frequentemente acompanha temperatura alta e baixa, tensão por salinidade ou secura, pode causar a desnaturação de de proteínas funcionais e estruturais. Como uma consequência, estas diversas tensões ambientais frequentemente ativam caminhos de sinalização celular e respostas celulares similares, tal como a produção de proteínas de tensão, super-regulação do acúmulo de anti-oxidantes, acúmulo dos solutos compatíveis e interrupção do desenvolvimento. O termo condições "sem tensão" como usado aqui são aquelas condições ambientais que permitem o desenvolvimento ótimo de planta de plantas. As pessoas habilitadas na técnica estão cientes de condições de solo normais e condições climáticas para uma dada localização. Plantas com condições de desenvolvimento ótimas, (desenvolvimento sob condições sem estresse) tipicamente, o rendimento na ordem crescente de preferência pelo menos 90 %, 87 %, 85 %, 83 %, 80 %, 77 % ou 75 % da produção média de tal planta em um dado ambiente. A produção média pode ser calculada com base na coleta e/ou na estação. As pessoas habilitadas na técnica estão cientes das produções de rendimento médio de uma lavoura.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições sem estresse ou sob condições de secura brandas com rendimento aumentado com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas que se desenvolvem sob condições sem estresse ou sob condições de secura brandas, cujo método compreende aumentar expressão em uma planta de um ácido

nucleico que codifica um polipeptídeo JMJC.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições de deficiência de nutriente, particularmente sob condições de deficiência de nitrogênio, rendimento aumentado com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas desenvolvidas sob condições de deficiência de nutriente, cujo método compreende aumentar expressão em uma planta de um ácido nucleico que codifica um polipeptídeo JMJC. A deficiência de nutriente pode resultar de uma perda ou de um excesso de nutrientes tais como nitrogênio, fosfatos e outros compostos contendo fósforo, potássio, cálcio, cádmio, magnésio, manganês, ferro e boro, entre outros.

A presente invenção abrange plantas ou partes destas (incluindo sementes) obteníveis pelos métodos de acordo com a presente invenção. As plantas ou partes destas compreende um ácido nucleico transgene que codifica um polipeptídeo JMJC como definido acima.

A invenção também fornece contruções genéticas e vetores para facilitar a introdução e/ou expressão em plantas de ácidos nucleicos que codificam polipeptídeos JMJC. A construção de genes pode ser inserida em vetores, que pode estar comercialmente disponível, adequada para a transformação em plantas e adequado para a expressão do gene de interesse nas células transformadas. A invenção também forence o uso da construção de gene como definido neste nos métodos da invenção.

Mais especificamente, a presente invenção fornece uma construção que compreende:

(a) um ácido nucleico que codifica um polipeptídeo JMJC como definido acima;

(b) uma ou mais sequências de controle capazes de conduzir a expressão da sequência de ácido nucleico de (a) e opcionalmente

(c) uma sequência de terminação de transcrição.

Preferivelmente, o ácido nucleico que codifica um polipeptídeo JMJC é como definido acima. O termo "sequência de controle" e "sequência de terminação" são como definidos neste.

As plantas são transformadas com um vetor que compreende qualquer um dos ácidos nucleicos descritos acima. O técnico habilitado estará bem ciente dos elementos genéticos que devem estar presentes no vetor na ordem para transformar, selecionar e propagar de maneira bem sucedida as células hospedeiras contendo a sequência de interesse. A sequência de interesse é operacionalmente ligada a uma ou mais sequências de controle (pelo menos a um promotor).

Vantajosamente, qualquer tipo de promotor, se natural ou sintético, pode ser usado para conduzir a expressão da sequência de ácido nucleico. Um promotor constitutivo é particularmente útil nos métodos. Ver a seção "definições" neste para definições dos vários tipos de promotores.

Deve estar claro que a aplicabilidade da presente invenção não está restrita ao polipeptídeo ácido nucleico que codifica JMJC representado pela SEQ ID N°: 73, nem é a aplicabilidade da invenção restrita à super-expressão de um polipeptídeo ácido nucleico que codifica JMJC quando conduzida por um promotor constitutivo.

O promotor constitutivo é preferivelmente um promotor GOS2, preferivelmente um promotor GOS2 de arroz. Ainda preferivelmente o promotor constitutivo é representado por uma sequência de ácido nucleico substancialmente similar à SEQ ID N°: 75, mais preferivelmente o promotor constitutivo é como representado pela SEQ ID N°: 75. Ver a seção "definições" neste para exemplos adicionais de promotores constitutivos.

Opcionalmente, uma ou mais sequências terminadoras podem ser usadas na construção introduzida em uma planta. Elementos reguladores adicionais podem incluir elementos de transcrição bem como de tradução.

Aqueles habilitados na técnica estarão cientes de sequências terminadoras e intensificadoras que podem ser adequados para o uso na realização da invenção. Uma sequência de íntron também pode ser adicionada à região 5' não traduzida (UTR) ou na sequência codificadora para aumentar a quantidade da mensagem madura que acumula-se no citosol, como descrito na seção de definições. Outras sequências de controle (além do promotor, intensificador, silenciador, sequências de íntron, regiões 3'UTR e/ou 5'UTR) podem ser elementos estabilziadores de proteína e/ou RNA. Tais sequências devem ser conhecidas ou podem ser facilmente obtidas por uma pessoa habilitada na técnica.

As construções genéticas da invenção ainda podem incluir uma origem de sequência de replicação que é requerida para a manutenção e/ou replicação em um tipo celular específico. Um exemplo ocorre quando a construção genética é requerida ser mantida em uma célula bacteriana como um elemento genético episssômico (por exemplo, molécula de plasmídeo ou cosmídeo). As origens preferidas de replicação incluem, mas não são limitadas a fl-ori e colEl.

Para a detecção da transferência bem sucedida das sequências de ácido nucleico como usado nos métodos da invenção e/ou seleção de plantas transgênicas que compreendem estes ácidos nucleicos, é vantajoso usar genes marcadores (ou genes repórteres). Portanto, a construção genética pode conter opcionalmente um gene de marcador selecionável. Marcadores selecionáveis são descritos em mais detalhes na seção "definições" neste.

É conhecido que na integração estável ou transitória de ácidos nucleicos em células de planta, apenas uma minoria das células absorve o DNA estranho e, se desejado, integra-o em seu genoma, dependendo do vetor de expressão usado e da técnica de transfecção usada. Para identificar e selecionar estes integrantes, um gene codificador para um marcador selecionável (tal como os descritos acima) é usualmente introduzido nas

células hospedeiras junto com o gene de interesse. Por exemplo,estes marcadores podem ser usados em mutantes em que estes genes não são funcionais, por exemplo, pela anulação pelos métodos convencionais. Além disso, moléculas de ácido nucleico que codificam um marcador selecionável podem ser introduzidas em uma célula hospedeira no mesmo vetor que compreende a sequência que codifica os polipeptídeos da invenção ou usados nos métodos da invenção ou ainda em um vetor separado. As células que foram estavelmente transfectadas com o ácido nucleico introduzido podem ser identificadas, por exemplo, pela seleção (por exemplo, células que integraram o marcador selecionável sobrevivem visto que as outras células morrem). Os genes marcadores podem ser removidos ou excisados da célula transgênica uma vez que estes não são mais necessários. As técnicas para a remoção do gene marcador são bem conhecidos na técnica, técnicas úteis são descritas acima na seção de definições.

A invenção também fornece um método para a produção de plantas transgênicas tendo traços intensificados relacionados com o rendimento com relação às plantas de controle, que compreende introdução e expressão em uma planta de qualquer ácido nucleico que codifica um polipeptídeo JMJC como definido neste acima.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo traços aumentados intensificados relacionados com o rendimento, particularmente rendimento aumentado (semente), cujo método compreende:

(i) introduzir e expressar em uma planta ou célula de planta um polipeptídeo ácido nucleico que codifica JMJC e

(ii) cultivar a célula de planta sob condições que promovam crescimento e desenvolvimento de planta.

O ácido nucleico de (i) pode ser qualquer um dos ácidos nucleicos capazes de codificar um polipeptídeo JMJC como definido neste.

O ácido nucleico pode ser introduzido diretamente em uma célula de planta ou na planta por si só (incluindo introdução em um tecido, órgão ou qualquer outra parte de uma planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é preferivelmente introduzido em uma planta por transformação. O termo "transformação" é descrito em mais detalhes na seção "definições" neste.

As células de planta geneticamente modificadas podem ser regeneradas por intermédio de todos os métodos com os quais o trabalhador habilitado está familiarizado. Os métodos adequados podem ser encontrados nas publicações mencionadas acima por S.D. Kung and R. Wu, Potrykus ou Hofgen e Willmitzer.

No geral após a transformação, as células de planta ou os agrupamentos de célula são selecionados quanto à presença de um ou mais marcadores que são codificados por genes expressíveis em plantas co-transferidas com o gene de interesse, seguindo que o material transformado é regenerado em uma planta total. Para selecionar as plantas transformadas, o material vegetal obtido na transformação é, como uma regra, submetido à condições seletivas de modo que as plantas transformadas possam ser distinguidas de plantas não transformadas. Por exemplo, as sementes obtidas da maneira descrita acima podem ser plantadas e, após um período de desenvolvimento, submetido a uma seleção adequada por pulverização. Uma possibilidade adicional consiste no desenvolvimento de sementes, se apropriado após a esterilização, em placas de ágar usando-se um agente de seleção adequado de modo que apenas as sementes transformadas possam se desenvolver em plantas. Alternativamente, as plantas transformadas são avaliadas quanto à presença de um marcador selecionável tal como os descritos acima.

Seguindo a transferência e a regeneração de DNA, as plantas putativamente transformadas também podem ser avaliadas, por exemplo,

usando-se análise de Southern, quanto à presença do gene de interesse, número de cópias e/ou organização genômica. Alternativa ou adicionalmente, os níveis de expressão do DNA recentemente introduzido pode ser monitorado usando-se Análise de Northern e/ou Western, ambas as técnicas sendo bem conhecidas pelas pessoas tendo habilidade comum na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de meios, tais como pela propagação clonal ou técnicas de criação clássicas. Por exemplo, uma planta transformada de primeira geração (ou Ti) podem ser tornadas independentes e transformantes de segunda geração homozigotos (ou T2) selecionados e as plantas T2 aidna podem ser propagadas através de técnicas de criação clássicas. Os organismos gerados podem tomar uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e de células não transformadas; transformantes clonais (por exemplo, todas as células transformadas contém o cassete de expressão); enxertos de tecidos transformados e não transformados (por exemplo, em plantas, um estoque de raiz transformado a um descendente não transformado).

A presente invenção claramente estende-se a qualquer célula de planta ou planta produzida por qualquer um dos métodos descritos neste e a todas as partes de plantas e seus propagules. A presente invenção estende-se ainda para abranger a progênie de uma célula, tecido, órgão ou planta total transformadas ou transfectadas primárias que foram produzidos por qualquer um dos métodos mencionados acima, o único requerimento sendo que a progênie apresente as mesmas características genotípicas e/ou fenotípicas como aqueles produzidos pelo precursor nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucleico isolado que codifica um polipeptídeo JMJC como definido neste acima. As células hospedeiras preferidas de acordo com a invenção são

células de planta. As plantas hospedeiras para o ácidos nucleicos ou para o vetor usadas no método de acordo com a invenção, o cassete de expressão ou construção ou vetor são, em princípio, vantajosamente todas as plantas, que são capazes de sintetizar os polipeptídeos usados no método inventivo.

Os métodos da invenção são vantajosamente aplicáveis a qualquer planta. Plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular, plantas monocotiledôneas e dicotiledôneas incluindo forragem e legumes de forragem, plantas ornamentais, lavouras de alimentos, árvores ou arbustos. De acordo com a forma de realização preferida da presente invenção, a planta é uma planta de cultivo. Exemplos de plantas de cultivo incluem soja, girassol, canola, alfafa, semente de colza, algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta monocotiledônea. Exemplos de plantas monocotiledôneas incluem cana de açúcar. Mais preferivelmente a planta é um cereal. Exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

A invenção também estende-se a partes colhíveis de uma planta tais como mas não limitadas a sementes, folhas, frutas, flores, caules, raízes, rizomas, tubérculos e bulbos. A invenção além disso diz respeito a produtos derivados, preferivelmente derivados diretamente de uma parte coletável de uma tal planta, tais como grânulos ou pós secos, óleo, gorura e ácidos graxos ou proteínas.

De acordo com uma característica preferida da invenção, a expressão modulada é a expressão aumentada. Métodos para aumentar a expressão de ácidos nucleicos ou genes ou produtos de gene, são bem documentados na técnica e os exemplos são fornecidos na seção de definições.

Como mencionado acima, um método preferido para modular (preferivelmente, aumentar) a expressão de um ácido nucleico que codifica

um polipeptídeo JMJC é pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo JMJC; entretanto os efeitos de realizar o método, isto é, intensificado os traços relacionados com o rendimento também pode ser atingido usando-se outras técnicas bem conhecidas, que incluem mas não limitado a rotulação de ativação de T-DNA, TILLING, recombinação homóloga. Uma descrição destas técnicas é fornecida na seção de definições.

A presente invenção também abrange o uso de ácidos nucleicos que codificam polipeptídeos JMJC como descrito neste e uso destes polipeptídeos JMJC na intensificação de qualquer um dos traços relacionados já mencionados com o rendimento em plantas.

Os ácidos nucleicos que codificam polipeptídeos JMJC descritos neste ou os polipeptídeos JMJC por si sós, podem encontrar uso nos programas de criação em que um marcador de DNA é identificado podendo estar geneticamente ligado a um gene que codifica o polipeptídeo JMJC. Os ácidos nucleicos/genes ou os polipeptídeos JMJC por si sós podem ser usados para definir um marcador molecular. Este DNA ou marcador de proteína podem ser então usados nos programas de criação para selecionar plantas tendo traços intensificados relacionados com o rendimento como definido neste acima nos métodos da invenção.

As variantes alélicas de um polipeptídeo ácido nucleico que codifica JMJC/gene também podem encontrar uso em programas de criação auxiliados por marcador. Tais programas de criação algumas vezes requerem a introdução de variação alélica pelo tratamento mutagênico das plantas, usando-se por exemplo Mutagênese por EMS; alternativamente, o programa pode iniciar com uma coleção de variantes alélicas da denominada origem "natural" não intencionalmente. A identificação de variantes alélicas então acontece, por exemplo, por PCR. Isto é seguido por uma etapa para a seleção de variantes alélicas superiores da sequência em questão e que dão

rendimento aumentado. Seleção é tipicamente realizada pelo monitoramento do desempenho de desenvolvimento de plantas contendo variantes alélicas da sequência em questão. Desempenho de desenvolvimento podem ser monitorados em uma estufa ou no campo. As etapas opcionais adicionais incluem o cruzamento de plantas em que a variante alélica foi identificada com uma planta. Isto pode ser usado, por exemplo, para fabricar uma combinação de características fenotípicas interessantes.

Ácidos nucleicos que codificam polipeptídeos JMJC também podem ser usados como sondas para mapear genética e fisicamente os genes que estes são uma parte de e como marcadores para traços ligados àqueles genes. Tal informação pode ser útil na criação de plantas a fim de desenvovler linhas com fenotipos desejados. Tal uso de ácidos nucleicos que codificam polipeptídeos JMJC requer apenas uma sequência de ácido nucleico de pelo menos 15 nucleotídeos de comprimento. The ácidos nucleicos que codificam polipeptídeos JMJC pode ser usado como marcadores de polimorfismo de comprimento de fragmento (RFLP). Southern blots (Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico de planta digerida por restrição pode ser sondada com o ácido nucleico que codifica JMJCs. Os padrões de formação de banda resultantes podem ser então submetidos a análises genéticas usando-se programas de computador, tais como MapMaker (Lander et al. (1987) Genomics 1: 174-181) a fim de construir um mampa genético. Além disso, os ácidos nucleicos podem ser usados para Southern blots de sonda contendo DNAs genômicos tratados com endonuclease de restrição de uma série de indivíduos que representam precursores e progênie de um cruzamento genético definido. Segregação dos Polimorfismos de DNA é observado e usado para calcular a posição do polipeptídeo ácido nucleico que codifica JMJC no mapa genético previamente obtido usando-se esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

A produção e o uso de sondas derivadas de gene vegetal para o uso no mapeamento genético são descritos em Bernatzky and Tanksley (1986) Plant Mol. Biol. Reporter 4: 37-41. Numerosas publicações descrevem o mapeamento genético de clones de cDNA específicos usando-se a metodologia resumida acima ou variações desta. Por exemplo, Populações intercruzamento F2, populações de retrocruzamento, populações aleatoriamente comparadas, linhas isogênicas próximas e outras séries de indivíduos pode ser usado para o mapeamento. Tais metodologias são bem conhecidas por aqueles habilitados na técnica.

As sondas de ácido nucleico também podem ser usadas para o mapeamento físico (isto é, colocação de sequências em mapas físicos; see Hoheisel et al. In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346 e referências citadas neste).

Em uma outra forma de realização, as sondas de ácido nucleico podem ser usadas no mapeamento de hibridização por fluorescência direta in situ (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora os métodos correntes de Mapeamento FISH favoreçam o uso de clones grandes (diversos kb a diversas centenas de kb; ver Laan et al. (1995) Genome Res. 5:13-20), melhoras na sensibilidade podem permitir o desempenho do mapeamento FISH usando-se sondas mais curtas.

Uma variedade de métodos com base na amplificação de ácido nucleico para o mapeamento genético e físico pode ser realizada usando-se o ácido nucleicos. os exemplos incluem a amplificação específica de alelo (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332), ligação específica de alelo (Landegren et al. (1988) Science 241:1077-1080), reações de extensão de nucleotídeo (Sokolov (1990) Nucleic Acid Res. 18:3671), mapeamento híbrido por radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e Mapeamento Happy (Dear and Cook (1989) Nucleic Acid Res.

17:6795-6807). Para estes métodos, a sequência de um ácido nucleico é usada para projetar e produzir pares de iniciador para o uso na reação de amplificação ou nas reações de extensão de iniciador. O projeto de tais iniciadores é bem conhecido por aqueles habilitados na técnica. Em métodos que utilizam Mapeamento genético com base em PCR, pode ser necesário identificar diferenças de sequência de DNA entre os precursores do mapeamento do cruzamento na região que corresponde à presente sequência de ácido nucleico. Isto, entretanto, não é necessário para os métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo traços intensificados relacionados com o rendimento, como descrito anteriormente. Estes traços também podem ser combinados com outros traços vantajosamente econômicos, tais como traços de intensificação de rendimento adicionais, tolerância a outras tensões bióticas ou abióticas, traços que modificam várias características arquitetônicas e/ou características bioquímicas e/ou fisiológicas.

III. Polipeptídeos CKI (Caseína quinase I)

Surpreendentemente, foi observado agora que modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo CKI dá plantas tendo traços intensificados relacionados com o rendimento com relação às plantas de controle. De acordo com uma primeira forma de realização, a presente invenção fornece um método para intensificar traços relacionados com o rendimento em plantas com relação às plantas de controle, que compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo CKI.

Um método preferido para modular (preferivelmente, aumentar) expressão de um ácido nucleico que codifica um polipeptídeo CKI é pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo CKI.

Qualquer referência a seguir a um "proteína útil nos métodos da invenção" é determinada entender um polipeptídeo CKI como definido neste. Qualquer referência a seguir a um "ácido nucleico útil nos métodos da invenção" é determinada entender um ácido nucleico capaz de codificar um tal polipeptídeo CKI. O ácido nucleico a ser introduzido em uma planta (e, portanto, útil na realização dos métodos da invenção) é qualquer ácido nucleico que codifica o tipo de proteína que será descrito agora, a seguir também denominado "CKI ácido nucleico" ou "gene CKI".

Um "polipeptídeo CKI" como definido neste refere-se às proteínas representadas pela SEQ ID N°: 174 e a homólogos (ortólogos e parálogos) destes. Preferivelmente, os homólogos da SEQ ID N°: 174 têm um domínio de caseína quinase.

O termo "domínio" e "motivo" é definido na seção "definições" neste. Os bancos de dados especialistas existem para a identificação de domínios, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-244), InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318), Prosite (Bucher and Bairoch (1994), Uma sintaxe de perfil generalizada para motivos de sequências biomoleculares e sua função em interpretação de sequência automática. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004)) ou Pfam (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002)). Uma série de ferramentas para a análise *in silico* de sequências de proteínas está disponível no servidor ExPASy proteomics (Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31:3784-3788(2003)). Os domínios eos motivos também podem ser identificados usando-se técnicas de rotina, tais como pelo

alinhamento de sequência.

Os métodos para o alinhamento de sequências para comparação são bem conhecidos na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. O GAP usa o algoritmo de Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar o alinhamento global (isto é medindo-se as sequências completas) de duas sequências que maximizam o número de equiparações e minimiza o número de fendas. O algoritmo BLAST (Altschul et at. (1990) J Mol Biol 215: 403-10) calcula a identidade de sequência percentual e realiza uma análise estatística da similaridade entre as duas sequências. O software para realizar análise BLAST está publicamente disponível através do National Centre for Biotechnology Information (NCBI). Os homólogos podem ser facilmente identificados usando-se, por exemplo, o algoritmo de alinhamento de sequência múltipla ClustalW (versão 1.83), com os parâmetros de alinhamento default e um método de registro em porcentagem. As Porcentagens globais de similaridade e de identidade também podem ser determinadas usando-se um dos métodos disponíveis no pacote de software MatGAT (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando-se proteína ou sequências de DNA.). a edição manual menor pode ser realizada para otimizar o alinhamento entre os motivos conservados, como deve estar evidente para uma pessoa habilitada na técnica. Além disso, em vez de usar sequências de comprimento total para a identificação de homólogos, domínios específicos também podem ser usados. Os valores de identidade de sequência podem ser determinados sobre o aminoácido inteiro ou sequência de aminoácido ou em domínios selecionados ou motivos conservados, usando-se os programas mencionados acima usando-se os parâmetros default.

Além disso, polipeptídeos CKI, tanto quanto SEQ ID N°: 174 e seus homólogos são considerados, as proteínas CKI úteis nos métodos nos

métodos da presente invenção tipicamente têm atividade de quinase. Os métodos para medir a atividade de quinase são bem conhecidos na técnica.

A presente invenção é ilustrada transformando-se as plantas com a sequência de ácido nucleico representado pela SEQ ID N°: 173, que codifica a sequência de polipeptídeo da SEQ ID N°: 174 respectivamente. Entretanto, o desempenho da invenção não está restrito a estas sequências; os métodos da invenção podem ser vantajosamente realizados usando-se qualquer ácido nucleico que codifica CKI ou polipeptídeo CKI como definido neste.

Os exemplos de ácidos nucleicos que codificam polipeptídeos CKI podem ser encontrados em bancos de dados conhecidos na técnica. tais ácidos nucleicos são úteis na realização dos métodos da invenção. Ortólogos e parálogos, os termos "ortólogos" e "parálogos" sendo como definido neste, podem ser facilmente identificados pela realização de uma denominada pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envovle submeter uma sequência em dúvida ao BLAST (por exemplo usando-se SEQ ID N°: 174) contra qualquer banco de dados de sequência, tal como os bancos de dados NCBI publicamente disponíveis. BLASTN or TBLASTX (usando-se os valores default padrão) são, no geral, usados quando começa-se a partir de uma sequência de nucleotídeo, and BLASTP or TBLASTN (usando-se os valores default padrão) quando começa-se a partir de uma sequência de proteína. Os resultados de BLAST podem ser opcionalmente filtrados. As sequências de comprimento total de cada um dos resultados filtrados ou resultados não filtrados são então submetidos ao BLAST novamente (segundo BLAST) contra sequências do organismo a partir do qual a sequência em questão é derivada (quando a sequência em questão é SEQ ID N°: 173 ou SEQ ID N°: 174, o segundo BLAST portanto, deve ser contra sequências de *Nicotiana tabacum*). Os resultados do primeiro e segundo BLASTs são então comparados. um parálogo é identificado se um

acerto de classificação alta do primeiro blast for da mesma espécie como a partir da qual a sequência em questão é derivada, submetendo-se ao BLAST novamente então resulta idealmente na sequência em questão entre os acertos mais altos; um ortólogo é identificado se um acerto de classificação alta no primeiro BLAST não for da mesma espécie como a partir da qual a sequência em questão é derivada e preferivelmente resulta no BLAST novamente na sequência em questão estando entre os acertos mais altos.

Os acertos de classificação alta são aqueles tendo um valor E baixo. Quanto mais baixo o valor E, mais significante o registro (ou, em outras palavras, menor a chance que o acerto seja encontrado por chance). Computação do valor E é bem conhecido na técnica. Além dos valores E, comparações são registradas pela identidade de porcentagem. Identidade de porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácidos nucleicos comparadas (ou polipeptídeo) em um comprimento particular. No caso de famílias grandes, ClustalW pode ser usado, seguido por uma árvore de união próxima, para ajudar a visualizar o agrupamento de genes relacionados e para identificar ortólogos e parálogos.

As variantes de ácido nucleico que codificam homólogos e derivados da SEQ ID N°: 174 também podem ser úteis na prática dos métodos da invenção, os termos "homólogo" e "derivado" sendo como definido neste. Também são úteis nos métodos da invenção os ácidos nucleicos que codificam homólogos e derivados de ortólogos or paralogues da SEQ ID N°: 174. Homólogos e derivados úteis nos métodos da presente invenção têm substancialmente a mesma atividade biológica e funcional como a proteína não modificada da qual estes são derivados.

As variantes de ácido nucleico adicionais úteis na prática dos métodos da invenção incluem porções de ácidos nucleicos que codificam polipeptídeos CKI, ácidos nucleicos que hibridizam-se a ácidos nucleicos que codificam polipeptídeos CKI, variantes de união de ácidos nucleicos que

codificam polipeptídeos CKI, variantes alélicas de ácidos nucleicos que codificam polipeptídeos CKI e variantes de ácidos nucleicos que codificam polipeptídeos CKI obtidos pelo embaralhamento de gene. Os termos sequência de hibridização, variante de união, variante alélica e embaralhamento de gene são como descritos neste.

Os ácidos nucleicos que codificam polipeptídeos CKI não necessitam ser ácidos nucleicos de comprimento total, visto que o desempenho dos métodos da invenção não contam com o uso de sequências de comprimento total de ácido nucleico. De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta uma porção da SEQ ID N°: 173 ou uma porção de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo da SEQ ID N°: 174.

Uma porção de um ácido nucleico pode ser preparada, por exemplo, realizando-se uma ou mais anulações no ácido nucleico. As porções podem ser usadas na forma isolada ou estas podem ser fundidas a outras sequências codificadoras (ou não codificadoras) a fim de, por exemplo, produzir uma proteína que combine diversas atividades. Quando fundido a outras sequências codificadoras, o polipeptídeo resultante produzido na tradução pode ser maior do que aquele predito para a porção de proteína.

As porções úteis métodos da invenção, codificam um polipeptídeo CKI como definido neste têm substancialmente a mesma atividade biológica como as sequências de aminoácido dada na SEQ ID N°: 174. Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos dada na SEQ ID N°: 173 ou é uma porção de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das sequências de aminoácido dada na SEQ ID N°: 173. Preferivelmente a porção is pelo menos 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1200, 1300, 1400 nucleotídeos de comprimento consecutivo, os nucleotídeos consecutivos

sendo da SEQ ID N°: 173 ou de um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 174. Mais preferivelmente a porção é uma porção do ácido nucleico da SEQ ID N°: 173.

Uma outra variante de ácido nucleico útil nos métodos da invenção é um ácido nucleico capaz de hibridizar, sob condições de estringência reduzida, preferivelmente sob condições estringentes, com um ácido nucleico que codifica um polipeptídeo CKI como definido neste ou com uma porção como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta, um ácido nucleico capable de hibridizar-se à SEQ ID N°: 173 ou que compreende introduzir e expressar em uma planta, um ácido nucleico capaz de hibridizar a um ácido nucleico que codifica um ortólogo, parálogo ou homólogo da SEQ ID N°: 173.

As sequências de hibridização úteis nos métodos da invenção codificam um polipeptídeo CKI como definido neste, tendo substancialmente a mesma atividade biológica as sequências de aminoácido dada na SEQ ID N°: 174. Preferivelmente, a sequência de hibridização é capaz de hibridizar à SEQ ID N°: 173 ou a uma porção de qualquer destas sequências, uma porção sendo como definida acima ou a sequência de hibridização é capaz de hibridizar a um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 174.

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante de união que codifica um polipeptídeo CKI como definido neste acima, a variante de união sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta a variante de união da SEQ ID N°: 173 ou a variante de união de um ácido nucleico que codifica um

ortólogo, parálogo ou homólogo da SEQ ID N°: 174.

Uma outra variante de ácido nucleico útil na realização dos métodos da invenção é uma variante alélica de um ácido nucleico que codifica um polipeptídeo CKI como definido neste acima, uma variante alélica sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta uma variante alélica da SEQ ID N°: 173 ou que compreende introduzir e expressar em uma planta uma variante alélica de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo das sequências de aminoácido representado pela SEQ ID N°: 174.

As variantes alélicas úteis nos métodos da presente invenção têm substancialmente a mesma atividade biológica como o polipeptídeo CKI da SEQ ID N°: 174. As variante alélicas existem em estado natural e abrangidas dentro dos métodos da presente invenção são o uso destes alelos naturais. Mistura de gene e a evolução direcionada podem ser usadas para gerar variantes de ácidos nucleicos que codificam polipeptídeos CKI como definido acima; o termo "mistura de gene" sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta a variante da SEQ ID N°: 173 ou que compreende introduzir e expressar em uma planta uma variante de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo da SEQ ID N°: 174, cujo ácido nucleico variante é obtido pelo embaralhamento de gene.

Além disso, as variantes de ácido nucleico também podem ser obtidas por mutagênese direcionada ao local. Os diversos métodos estão disponíveis para atingir mutagênese direcionada ao local, o mais comum

sendo métodos com base em PCR (Current Protocols in Molecular Biology. Wiley Eds.).

Os ácidos nucleicos que codificam polipeptídeos CKI podem ser derivados de qualquer fonte natural ou artificial. O ácido nucleico pode ser modificado a partir de sua forma natural na composição e/ou ambiente genômico através da manipulação humana deliberada. Preferivelmente, o ácido nucleico que codifica o polipeptídeo CKI é de uma planta. No caso SEQ ID N°: 173, os ácidos nucleicos que codificam polipeptídeo CKI é preferivelmente de uma planta dicotiledônea, mais preferivelmente da família Solanaceae, mais preferivelmente o ácido nucleico é de *Nicotiana tabacum.*

O desempenho dos métodos da invenção dá plantas tendo traços intensificados relacionados com o rendimento. Em particular, desempenho dos métodos da invenção dá plantas tendo vigor precoce aumentado e rendimento aumentado, especialmente biomassa aumentada e rendimento aumentado de semente com relação às plantas de controle. Os termos "rendimento" e "rendimento de semente" são descritos em mais detalhes na seção "definições" neste.

Referência neste a traços intensificados relacionados com o rendimento é determinada entender um aumento em vigor precoce e/ou na biomassa (peso) de uma ou mais partes de uma planta, que pode incluir partes (coletáveis) acima do solo e/ou partes (coletáveis) abaixo do solo. Em particular, tais partes colhíveis são biomassa e/ou sementes e desempenho dos métodos da invenção resulta em plantas tendo vigor precoce aumentado, biomassa e/ou rendimento de semente com relação à vigor precoce, biomassa ou rendimento de semente de plantas de controle.

Tomando o milho como um exemplo, um aumento no rendimento pode ser manifestado as um ou mais do seguinte: aumento no número de plantas estabelecido por hectare ou acre, um aumento no número de espigas por planta, um aumento no número de fileiras, número de sementes

por fileira, peso de semente, peso de semente milhar, comprimento/diâmetro de espiga, aumento na razão de enchimento de semente (que é o número de sementes enchidas dividido pelo número total de sementes e multiplicado por 100), entre outros. Tomando-se o arroz como um exemplo, um aumento no rendimento pode manifestar-se por si só como um aumento em um ou mais do seguinte: número de plantas por hectare ou acre, número de panículas por planta, número de espigas por panícula, número de flores (flósculos) por panícula (que é expressado como uma razão do número de sementes enchidas no número de panícula primárias), aumento na razão de enchimento de semente (que é o número de sementes enchidas dividido pelo número total de sementes e multiplicado por 100), aumento em peso de semente milhar, entre outros.

A presente invenção fornece um método para aumentar o rendimento, especialmente biomassa e/ou rendimento de semente de plantas, com relação às plantas de controle, cujo método compreende a expressão modular, preferivelmente aumentar a expressão, em uma planta de um ácido nucleico que codifica um polipeptídeo CKI como definido neste.

Visto que as plantas transgênicas de acordo com a presente invenção têm rendimento aumentado, é provável que estas plantas apresentem uma taxa de desenvolvimento aumentada (durante pelo menos parte do seu ciclo de vida), com relação à taxa de desenvolvimento de plantas de controle em um estágio correspondente em seu ciclo de vida.

A taxa de desenvolvimento aumentada pode ser específica para uma ou mais partes de uma planta (incluindo sementes) ou pode ser substancialmente por toda a planta. Plantas tendo uma taxa de desenvolvimento aumentada pode ter um ciclo de vida mais curto. O ciclo de vida de uma planta pode ser tomado significar o tempo necessário para o desenvolvimento de uma semente madura seca até o estágio onde a planta produziu sementes maduras secas, similar ao material de partida. Este ciclo de

vida pode ser influenciado por fatores tais como vigor precoce, taxa de desenvolvimento, índice de verdor, tempo de florescimento e velocidade de maturação de semente. O aumento em taxa de desenvolvimento pode acontecer em um ou mais estágios no ciclo de vida de uma planta ou substancialmente durante o ciclo de vida de planta total. A taxa de desenvolvimento aumentada durante os estágios precoces no ciclo de vida de uma planta pode refletir o vigor intensificado. O aumento na taxa de desenvolvimento pode alterar o ciclo de coleta de uma planta permitindo que as plantas sejam semeadas depois e/ou coletadas mais cedo do que seria de outra maneira possível (um efeito similar pode ser obtido com tempo de florescimento precoce). Se a taxa de desenvolvimento for suficientemente aumentada, isto pode permitir a semeadura adicional de sementes das mesmas espécies vegetais (por exemplo semeadura e coleta de plantas de arroz seguido pela semeadura e coleta de outras plantas de arroz tudo dentro de um período de desenvolvimento convencional). Similarmente, se a taxa de desenvolvimento for suficientemente aumentada, isto pode permitir a semeadura adicional de sementes de espécies de plantas diferentes (por exemplo a semeadura e a coleta de plantas de milho seguido por, por exemplo, a semeadura e a coleta opcional de soja, batata ou qualquer outro vegetal adequado). Os períodos de coleta adicionais a partir do mesmo estoque de raiz no caso de algumas plantas de cultivo também pode ser possível. Alteração do ciclo de coleta de uma planta pode levar a um aumento na produção de biomassa anual por acre (devido a um aumento no número de vezes (diz-se em um ano) que qualquer planta particular pode ser desenvolvida e coletada). Um aumento na taxa de desenvolvimento também pode permitir o cultivo de plantas transgênicas em uma área geográfica mais ampla do que suas contrapartes do tipo selvagem, visto que as limitações territoriais para o desenvolvimento de uma lavoura são frequentemente determinados pelas condições ambientais adversas no período de plantio

(estação precoce) ou no período de coleta (estação posterior). Tais condições adversas podem ser evitadas se o ciclo de coleta for encurtado. A taxa de desenvolvimento pode ser determinada derivando-se vários parâmetros de curvas de desenvolvimento, tais parâmetros podem ser: T-Mid (o tempo levado pelas plantas para atingir 50 % de seu tamanho máximo) e T-90 (tempo levado pelas plantas para atingir 90 % de seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, desempenho dos métodos da invenção dá plantas tendo uma taxa de desenvolvimento aumentada com relação às plantas de controle. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a taxa de desenvolvimento de plantas, cujo método compreende a expressão modular, preferivelmente aumentar a expressão, em uma planta de um ácido nucleico que codifica um polipeptídeo CKI como definido neste. Em uma forma de realização particular, desempenho dos métodos da presente invenção dá plantas com vigor precoce aumentado.

Um aumento no rendimento e/ou taxa de desenvolvimento ocorre se a planta estiver sob condições sem estresse ou se a planta for exposta a várias tensões em comparação com as plantas de controle. Plantas tipicamente respondem à exposição à tensão desenvolvendo-se mais lentamente. Em condições de tensão grave, a planta pode ainda parar de se desenvolver completamente. A tensão branda por outro lado é definida neste como sendo qualquer tensão à qual uma planta é exposta que não resulta na planta parando seu desenvolvimento completamente sem a capacidade de retomar o crescimento. Tensão branda no sentido da invenção da invenção leva a uma redução no desenvolvimento das plantas submetidas à tensão de menos que 40 %, 35 % ou 30 %, preferivelmente menor do que 25 %, 20 % ou 15 %, mais preferivelmente menor do que 14 %, 13 %, 12 %, 11 % ou 10 % ou menor em comparação com a planta de controle sob condições sem

estresse. Devido aos avanços nas práticas agrícolas (irrigação, fertilização e tratamentos com pesticidas) tensões graves não são frequentemente encontradas em plantas de cultivo cultivadas. Como uma consequência, o desenvolvimento comprometido induzido pela tensão branda é frequentemente uma característica indesejada para a agricultura. As tensões brandas são as tensões (ambientais) bióticas e/ou abióticas diárias às quais a planta é exposta. As tensões abióticas podem ocorrer devido à secura ou excesso de água, tensão anaeróbica, tensão ao sal, toxicidade química, tensão oxidativa e calor, frio ou temperaturas congelantes. A tensão abiótica pode ser uma tensão osmótica causada por uma tensão a água (particularmente devido à secura), tensão ao sal, tensão oxidativa ou uma tensão iônica. Tensões bióticas são tipicamente aquelas tensões causadas por patogênio, tal como bactérias, vírus, fungos e insetos.

Em particular, os métodos da presente invenção podem ser realizados sob condições sem estresse ou sob condições de secura branda para dar plantas tendo rendimento aumentado com relação às plantas de controle. Como relatado em Wang et al. (Plant (2003) 218: 1-14), tensão abiótica leva a uma série de mudanças morfológicas, fisiológica, bioquímicas e moleculares que afetam adversamente o desenvolvimento e produtividade da planta. Secura, salinidade, temperaturas extremas e tensão oxidativa são conhecidas serem interconectadas e podem induzir o desenvolvimento e o dano celular através de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve um grau particularmente alto de "cross talk" tensão por secura e tensão por salinidade alta. Por exemplo, seca e/ou desalinização são manifestadas primariamente como tensão osmótica, resultando na interrupção da homeostase e a distribuição de íon na célula. Tensão oxidativa, que frequentemente acompanha temperatura alta e baixa, tensão por salinidade ou secura, pode causar a desnaturação de de proteínas funcionais e estruturais. Como uma consequência, estas diversas tensões ambientais frequentemente

ativam caminhos de sinalização celular e respostas celulares similares, tal como a produção de proteínas de tensão, super-regulação do acúmulo de anti-oxidantes, acúmulo dos solutos compatíveis e interrupção do desenvolvimento. O termo condições "sem tensão" como usado aqui são aquelas condições ambientais que permitem o desenvolvimento ótimo de planta de plantas. As pessoas habilitadas na técnica estão cientes de condições de solo normais e condições climáticas para uma dada localização. Plantas com condições de desenvolvimento ótimas, (desenvolvimento sob condições sem estresse) tipicamente, o rendimento na ordem crescente de preferência pelo menos 90 %, 87 %, 85 %, 83 %, 80 %, 77 % ou 75 % da produção média de tal planta em um dado ambiente. A produção média pode ser calculada com base na coleta e/ou na estação. As pessoas habilitadas na técnica estão cientes das produções de rendimento médio de uma lavoura.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições sem estresse ou sob condições de secura brandas rendimento aumentado e/ou vigor precoce aumentado, com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento e/ou vigor precoce em plantas que se desenvolvem sob condições sem estresse ou sob condições de secura brandas, cujo método compreende aumentar expressão em uma planta de um ácido nucleico que codifica um polipeptídeo CKI.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições de deficiência de nutriente, particularmente sob condições de deficiência de nitrogênio, rendimento aumentado com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas desenvolvidas sob condições de deficiência de nutriente, cujo método compreende aumentar expressão em uma planta de um

ácido nucleico que codifica um polipeptídeo CKI. Deficiência de nutriente pode resultar from a lack of nutrients tais como nitrogênio, fosfatos e outros compostos contendo fósforo, potássio, cálcio, cádmio, magnésio, manganês, ferro e boro, entre outros.

A presente invenção abrange plantas ou partes destas (incluindo sementes) obteníveis pelos métodos de acordo com a presente invenção. As plantas ou partes destas compreende um ácido nucleico transgene que codifica um polipeptídeo CKI como definido acima.

A invenção também fornece contruções genéticas e vetores para facilitar a introdução e/ou expressão em plantas de ácidos nucleicos que codificam polipeptídeos CKI. A construção de genes pode ser inserida em vetores, que pode estar comercialmente disponível, adequada para a transformação em plantas e adequado para a expressão do gene de interesse nas células transformadas. A invenção também forence o uso da construção de gene como definido neste nos métodos da invenção.

Mais especificamente, a presente invenção fornece uma construção que compreende:

(a) um ácido nucleico que codifica um polipeptídeo CKI como definido acima;

(b) uma ou mais sequências de controle capazes de conduzir a expressão da sequência de ácido nucleico de (a) e opcionalmente

(c) uma sequência de terminação de transcrição.

Preferivelmente, o ácido nucleico que codifica um polipeptídeo CKI é como definido acima. O termo "sequência de controle" e "sequência de terminação" são como definidos neste.

As plantas são transformadas com um vetor que compreende qualquer um dos ácidos nucleicos descritos acima. O técnico habilitado estará bem ciente dos elementos genéticos que devem estar presentes no vetor na ordem para transformar, selecionar e propagar de maneira bem sucedida as

células hospedeiras contendo a sequência de interesse. A sequência de interesse é operacionalmente ligada a uma ou mais sequências de controle (pelo menos a um promotor).

Vantajosamente, qualquer tipo de promotor, se natural ou sintético, pode ser usado para conduzir a expressão da sequência de ácido nucleico. Um promotor específico de semente é particularmente útil nos métodos da invenção, preferivelmente, o promotor é um promotor específico de embrião. Ver a seção "definições" neste para definições dos vários tipos de promotores. Também úteis nos métodos da invenção is a promotor constitutivo.

Deve estar claro que a aplicabilidade da presente invenção não está restrita ao ácido nucleico que codifica o polipeptídeo CKI representado pela SEQ ID N°: 173, nem é a aplicabilidade da invenção restrita à super-expressão de um ácido nucleico que codifica o polipeptídeo CKI quando conduzida por um promotor específico de semente ou quando conduzida por um promotor constitutivo.

O promotor específico de semente é preferivelmente um promotor WSI18, preferivelmente um promotor WSI18 de arroz. Ainda preferivelmente o promotor específico de semente é representado por uma sequência de ácido nucleico substancialmente similar à SEQ ID N°: 175, mais preferivelmente o promotor específico de semente é como representado pela SEQ ID N°: 175. Ver a seção "definições" neste para exemplos adicionais de promotor específico de sementes.

Opcionalmente, uma ou mais sequências terminadoras podem ser usadas na construção introduzida em uma planta. Elementos reguladores adicionais podem incluir elementos de transcrição bem como de tradução. Aqueles habilitados na técnica estarão cientes de sequências terminadoras e intensificadoras que podem ser adequados para o uso na realização da invenção. Uma sequência de íntron também pode ser adicionada à região 5'

não traduzida (UTR) ou na sequência codificadora para aumentar a quantidade da mensagem madura que acumula-se no citosol, como descrito na seção de definições. Outras sequências de controle (além do promotor, intensificador, silenciador, sequências de íntron, regiões 3'UTR e/ou 5'UTR) podem ser elementos estabilziadores de proteína e/ou RNA. Tais sequências devem ser conhecidas ou podem ser facilmente obtidas por uma pessoa habilitada na técnica.

As construções genéticas da invenção ainda podem incluir uma origem de sequência de replicação que é requerida para a manutenção e/ou replicação em um tipo celular específico. Um exemplo ocorre quando a construção genética é requerida ser mantida em uma célula bacteriana como um elemento genético epissômico (por exemplo, molécula de plasmídeo ou cosmídeo). As origens preferidas de replicação incluem, mas não são limitadas a fl-ori e colEl.

Para a detecção da transferência bem sucedida das sequências de ácido nucleico como usado nos métodos da invenção e/ou seleção de plantas transgênicas que compreendem estes ácidos nucleicos, é vantajoso usar genes marcadores (ou genes repórteres). Portanto, a construção genética pode conter opcionalmente um gene de marcador selecionável. Os marcadores selecionáveis são descritos em mais detalhes na seção "definições" neste. Os genes marcadores podem ser removidos ou excisados da célula transgênica uma vez que estes não são mais necessários. As técnicas para a remoção de marcador são bem conhecidas na técnica, as técnicas úteis são descritas acima na seção de definições.

A invenção também fornece um método para a produção de plantas transgênicas tendo traços intensificados relacionados com o rendimento com relação às plantas de controle, que compreende introdução e expressão em uma planta de qualquer ácido nucleico que codifica um polipeptídeo CKI como definido neste acima.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo traços aumentados intensificados relacionados com o rendimento, particularmente vigor precoce aumentado e/ou rendimento aumentado, cujo método compreende:

(i) introduzir e expressar em uma planta ou célula de planta um polipeptídeo ácido nucleico que codifica CKI e

(ii) cultivar a célula de planta sob condições que promovam crescimento e desenvolvimento de planta.

O ácido nucleico de (i) pode ser qualquer um dos ácidos nucleicos capazes de codificar um polipeptídeo CKI como definido neste.

O ácido nucleico pode ser introduzido diretamente em uma célula de planta ou na planta por si só (incluindo introdução em um tecido, órgão ou qualquer outra parte de uma planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é preferivelmente introduzido em uma planta por transformação. O termo "transformação" é descrito em mais detalhes na seção "definições" neste.

As células de planta geneticamente modificadas podem ser regeneradas por intermédio de todos os métodos com os quais o trabalhador habilitado está familiarizado. Os métodos adequados podem ser encontrados nas publicações mencionadas acima por S.D. Kung and R. Wu, Potrykus ou Hofgen e Willmitzer.

No geral após a transformação, as células de planta ou os agrupamentos de célula são selecionados quanto à presença de um ou mais marcadores que são codificados por genes expressíveis em plantas co-transferido com o gene de interesse, seguindo que o material transformado é regenerado em uma planta total. Para selecionar as plantas transformadas, o material vegetal obtido na transformação é, como uma regra, submetido à condições seletivas de modo que as plantas transformadas possam ser distinguidas das plantas não transformadas. Por exemplo, as sementes obtidas

da maneira descrita acima podem ser plantadas e, após um período de desenvolvimento, submetido a uma seleção adequada por pulverização. Uma possibilidade adicional consiste no desenvolvimento de sementes, se apropriado após a esterilização, em placas de ágar usando-se um agente de seleção adequado de modo que apenas as sementes transformadas possam se desenvolver em plantas. Alternativamente, as plantas transformadas são avaliadas quanto à presença de um marcador selecionável tal como os descritos acima.

Seguindo a transferência e a regeneração de DNA, as plantas putativamente transformadas também podem ser avaliadas, por exemplo, usando-se análise de Southern, quanto à presença do gene de interesse, número de cópias e/ou organização genômica. Alternativa ou adicionalmente, os níveis de expressão do DNA recentemente introduzido pode ser monitorado usando-se Análise de Northern e/ou Western, ambas as técnicas sendo bem conhecidas pelas pessoas tendo habilidade comum na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de meios, tais como pela propagação clonal ou técnicas de criação clássicas. Por exemplo, uma planta transformada de primeira geração (ou T1) podem ser tornadas independentes e transformantes de segunda geração homozigotos (ou T2) selecionados e as plantas T2 ainda podem ser propagadas através de técnicas de criação clássicas. Os organismos gerados podem tomar uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e de células não transformadas; transformantes clonais (por exemplo, todas as células transformadas contém o cassete de expressão); enxertos de tecidos transformados e não transformados (por exemplo, em plantas, um estoque de raiz transformado a um descendente não transformado).

A presente invenção claramente estende-se a qualquer célula de planta ou planta produzida por qualquer um dos métodos descritos neste e

a todas as partes de plantas e seus propagules. A presente invenção estende-se ainda para abranger a progênie de uma célula, tecido, órgão ou planta total transformadas ou transfectadas primárias que foram produzidos por qualquer um dos métodos mencionados acima, o único requerimento sendo que a progênie apresente as mesmas características genotípicas e/ou fenotípicas como aqueles produzidos pelo precursor nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucleico isolado que codifica um polipeptídeo CKI como definido neste acima. As células hospedeiras preferidas de acordo com a invenção são células de planta. As plantas hospedeiras para o ácidos nucleicos ou para o vetor usadas no método de acordo com a invenção, o cassete de expressão ou construção ou vetor são, em princípio, vantajosamente todas as plantas, que são capazes de sintetizar os polipeptídeos usados no método inventivo.

Os métodos da invenção são vantajosamente aplicáveis a qualquer planta. As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular, plantas monocotiledôneas e dicotiledôneas incluindo forragem e legumes de forragem, plantas ornamentais, lavouras de alimentos, árvores ou arbustos. De acordo com a forma de realização preferida da presente invenção, a planta é uma planta de cultivo. Os exemplos de plantas de cultivo incluem soja, girassol, canola, alfafa, semente de colza, algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta monocotiledônea. Os exemplos de plantas monocotiledôneas incluem cana de açúcar. Mais preferivelmente a planta é um cereal. Os exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

A invenção também estende-se a partes colhíveis de uma planta tais como mas não limitadas a sementes, folhas, frutas, flores, caules, raízes, rizomas, tubérculos e bulbos. A invenção além disso diz respeito a

produtos derivados, preferivelmente derivados diretamente de uma parte coletável de uma tal planta, tais como grânulos ou pós secos, óleo, gorura e ácidos graxos ou proteínas.

De acordo com uma característica preferida da invenção, a expressão modulada é a expressão aumentada. Métodos para aumentar a expressão de ácidos nucleicos ou genes ou produtos de gene, são bem documentados na técnica e os exemplos são fornecidos na seção de definições.

Como mencionado acima, um método preferido para modular (preferivelmente, aumentar) a expressão de um ácido nucleico que codifica um polipeptídeo CKI é pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo CKI; entretanto os efeitos de realizar o método, isto é, intensificado os traços relacionados com o rendimento também pode ser atingido usando-se outras técnicas bem conhecidas, que incluem mas não limitado a rotulação de ativação de T-DNA, TILLING, recombinação homóloga. Uma descrição destas técnicas é fornecida na seção de definições.

A presente invenção também abrange o uso de ácidos nucleicos que codificam polipeptídeos CKI como descrito neste e uso destes polipeptídeos CKI na intensificação de qualquer um dos traços relacionados já mencionados com o rendimento em plantas.

Os ácidos nucleicos que codificam polipeptídeo CKI descritos neste ou os polipeptídeos CKI por si sós, podem encontrar uso nos programas de criação em que um marcador de DNA é identificado podendo estar geneticamente ligado a um gene que codifica o polipeptídeo CKI. O ácido nucleicos/genes ou os polipeptídeos CKI por si sós podem ser usados para definir um marcador molecular. Este amrcador de DNA ou proteína podem ser então usados nos programas de criação para selecionar plantas tendo traços intensificados relacionados com o rendimento como definido neste

acima nos métodos da invenção.

As variantes alélicas de um ácido nucleico/gene que codifica o polipeptídeo CKI também podem encontrar uso em programas de criação auxiliados por marcador. Tais programas de criação algumas vezes requerem a introdução de variação alélica pelo tratamento mutagênico das plantas, usando-se por exemplo mutagênese por EMS; alternativamente, o programa pode iniciar com uma coleção de variantes alélicas da denominada origem "natural" não intencionalmente. A identificação de variantes alélicas então acontece, por exemplo, por PCR. Isto é seguido por uma etapa para a seleção de variantes alélicas superiores da sequência em questão e que dão rendimento aumentado. A seleção é tipicamente realizada pelo monitoramento do desempenho de desenvolvimento de plantas contendo variantes alélicas da sequência em questão. O desempenho de desenvolvimento pode ser monitorado em uma estufa ou no campo. As etapas opcionais adicionais incluem o cruzamento de plantas em que a variante alélica foi identificada com uma planta. Isto pode ser usado, por exemplo, para fabricar uma combinação de características fenotípicas interessantes.

Os ácidos nucleicos que codificam polipeptídeos CKI também podem ser usados como sondas para mapear genética e fisicamente os genes que estes são uma parte de e como marcadores para traços ligados àqueles genes. Tal informação pode ser útil na criação de plantas a fim de desenvovler linhas com fenotipos desejados. Tal uso de ácidos nucleicos que codificam polipeptídeo CKI requer apenas uma sequência de ácido nucleico de pelo menos 15 nucleotídeos de comprimento. Os ácidos nucleicos que codificam polipeptídeo CKI codificam ácidos nucleicos pode ser usado como marcadores de polimorfismo de comprimento de fragmento (RFLP). Southern blots (Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico de planta digerida por restrição pode ser sondada com os ácidos nucleicos que codificam CKI. Os padrões de

formação de banda resultantes podem ser então submetidos a análises genéticas usando-se programas de computador, tais como MapMaker (Lander et al. (1987) Genomics 1: 174-181) a fim de construir um mampa genético. Além disso, os ácidos nucleicos podem ser usados para Southern blots de sonda contendo DNAs genômicos tratados com endonuclease de restrição de uma série de indivíduos que representam precursores e progênie de um cruzamento genético definido. Segregação dos Polimorfismos de DNA é observado e usado para calcular a posição do ácido nucleico que codifica o polipeptídeo CKI no mapa genético previamente obtido usando-se esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

A produção e o uso de sondas derivadas de gene vegetal para o uso no mapeamento genético são descritos em Bernatzky and Tanksley (1986) Plant Mot. Biol. Reporter 4: 37-41. Numerosas publicações descrevem o mapeamento genético de clones de cDNA específicos usando-se a metodologia resumida acima ou variações desta. Por exemplo, Populações intercruzamento F2, populações de retrocruzamento, populações aleatoriamente comparadas, linhas isogênicas próximas e outras séries de indivíduos pode ser usado para o mapeamento. Tais metodologias são bem conhecidas por aqueles habilitados na técnica.

As sondas de ácido nucleico também podem ser usadas para o mapeamento físico (isto é, colocação de sequências em mapas físicos; see Hoheisel et al. In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346 e referências citadas neste).

Em uma outra forma de realização, as sondas de ácido nucleico pode ser usadas no mapeamento de hibridização por fluorescência direta in situ (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora os métodos correntes de Mapeamento FISH favoreçam o uso de clones grandes (diversos kb a diversas centenas de kb; ver Laan et al. (1995) Genome Res. 5:13-20), melhoras na sensibilidade podem permitir o desempenho do

mapeamento FISH usando-se sondas mais curtas.

Uma variedade de métodos com base na amplificação de ácido nucleico para o mapeamento genético e físico pode ser realizada usando-se o ácido nucleicos. Os exemplos incluem a amplificação específica de alelo (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield et at. (1993) Genomics 16:325-332), ligação específica de alelo (Landegren et al. (1988) Science 241:1077-1080), reações de extensão de nucleotídeo (Sokolov (1990) Nucleic Acid Res. 18:3671), mapeamento híbrido por radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e Mapeamento Happy (Dear and Cook (1989) Nucleic Acid Res. 17:6795-6807). Para estes métodos, a sequência de um ácido nucleico é usada para projetar e produzir pares de iniciador para o uso na reação de amplificação ou nas reações de extensão de iniciador. O projeto de tais iniciadores é bem conhecido por aqueles habilitados na técnica. Em métodos que utilizam mapeamento genético com base em PCR, pode ser necessário identificar diferenças de sequência de DNA entre os precursores do mapeamento do cruzamento na região que corresponde à presente sequência de ácido nucleico. Isto, entretanto, não é necessário para os métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo traços intensificados relacionados com o rendimento, como descrito anteriormente. Estes traços também podem ser combinados com outros traços vantajosamente econômicos, tais como traços de intensificação de rendimento adicionais, tolerância a outras tensões bióticas ou abióticas, traços que modificam várias características arquitetônicas e/ou características bioquímicas e/ou fisiológicas.

IV. Polipeptídeo de prolongamento-homeodomínio de homeodomíno de planta (PHDf-HD)

Surpreendentemente, foi observado agora que modular,

preferivelmente aumentar, expressão em uma planta de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD dá plantas tendo traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, com relação às plantas de controle. De acordo com uma primeira forma de realização, a presente invenção fornece um método para intensificar traços relacionados com o rendimento, preferivelmente enhancing traços de sementes relacionados com o rendimento, em plantas com relação às plantas de controle, que compreende modular, preferivelmente aumentar, expressão em uma planta de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD.

Um método preferido para modular, preferivelmente aumentar, a expressão de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD ocorre pela introdução e expressão em uma planta de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD.

Qualquer referência a seguir a um "proteína útil nos métodos da invenção" é determinada entender um polipeptídeo PHDf-HD como definido neste. Qualquer referência a seguir a uma "sequência de ácido nucleico útil nos métodos da invenção" é determinada entender uma sequência de ácido nucleico capaz de codificar um tal polipeptídeo PHDf-HD. A sequência de ácido nucleico a ser introduzido em uma planta (e, portanto, útil na realização dos métodos da invenção) é qualquer sequência de ácido nucleico que codifica o tipo de polipeptídeo, que será descrito agora, a seguir também denominado "sequência de ácido nucleico PHDf-HD" ou "gene PHDf-HD".

Um "polipeptídeo PHDf-HD" como definido neste refere-se a qualquer polipeptídeo que compreende: (i) um domínio tendo pelo menos 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais identidades de sequência de aminoácido

a um domínio zíper de leucina/prolongamento de homeodomínio de planta (ZIP/PHDf) como representado pela SEQ ID N°: 233; e (ii) um domínio tendo pelo menos 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais identidades de sequência de aminoácido a um homeodomínio (HD) como representado pela SEQ ID N°: 234.

Alternativa ou adicionalmente, um "polipeptídeo PHDf-HD" como definido neste refere-se a qualquer polipeptídeo que compreende: (i) um domínio PHD como representado pela PFAM00628; e (ii) um HD como representado pela PFAM00046.

Alternativa ou adicionalmente, um "polipeptídeo PHDf-HD" como definido neste refere-se a qualquer sequência de polipeptídeo que quando usada na construção de uma árvore filogenética HD, tal como o descrito na Figura 13, agrupa-se com os grupo PHDf-HD de polipeptídeos que compreende a sequência de polipeptídeo como representado pela SEQ ID N°: 180, em vez de qualquer outro grupo HD.

Alternativa ou adicionalmente, um "polipeptídeo PHDf-HD" como definido neste refere-se a qualquer polipeptídeo tendo na ordem crescente de preferência pelo menos 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais identidades de sequência de aminoácido ao polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180 ou a qualquer uma das sequências de polipeptídeos dadas na Tabela D1 neste.

O termo "domínio" e "motivo" é definido na seção "definições" neste. Os bancos de dados especialistas existem para a identificação de domínios, por exemplo, SMART (Schultz et at. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et at. (2002) Nucleic Acids Res 30, 242-244, InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318, Prosite (Bucher and Bairoch (1994), Uma sintaxe de perfil generalizada para

motivos de sequências biomoleculares e sua função em interpretação de sequência automática. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hula et al., Nucl. Acids. Res. 32:D134-D137, (2004) ou Pfam (Bateman et at., Nucleic Acids Research 30(1): 276-280 (2002). Uma série de ferramentas para a análise *in silico* de sequências de proteínas está disponível no servidor ExPASy proteomics (Swiss Institute of Bioinformatics (Gasteiger et at., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31:3784-3788(2003)). Os domínios também podem ser identificados usando-se técnicas de rotina, tais como pelo alinhamento de sequência. A análise da sequência de polipeptídeo da SEQ ID N°: 180 é apresentada abaixo nos Exemplos 30 e 32 neste. Por exemplo, um polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180 compreende um dedo de PHD com uma entrada Pfam PF00628 e um HD com a entrada Pfam PF00046.

Os métodos para o alinhamento de sequências para comparação são bem conhecidos na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. O GAP usa o algoritmo de Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar o alinhamento global (isto é medindo-se as sequências completas) de duas sequências que maximizam o número de equiparações e minimiza o número de fendas. O algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-10) calcula a identidade de sequência percentual e realiza uma análise estatística da similaridade entre as duas sequências. O software para realizar análise BLAST está publicamente disponível através do National Centre for Biotechnology Information (NCBI). Os homólogos podem ser facilmente identificados usando-se, por exemplo, o algoritmo de alinhamento de sequência múltipla ClustalW (versão 1.83), com os parâmetros de

alinhamento default e um método de registro em porcentagem. As porcentagens globais de similaridade e de identidade também podem ser determinados usando-se um dos métodos disponíveis no pacote de software MatGAT (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando-se proteína ou sequências de DNA.). edição manual menor pode ser realizada para otimizar o alinhamento entre os motivos conservados, como deve estar evidente para uma pessoa habilitada na técnica. Além disso, em vez de usar sequências de comprimento total para a identificação de homólogos, domínios específicos também podem ser usados. Os valores de identidade de sequência podem ser determinados sobre a sequência de aminoácido inteira ou a sequência de polipeptídeo ou em domínios selecionados ou motivos conservados, usando-se os programas mencionados acima usando-se os parâmetros default. Example 31 herein describes in Table D2 the identidade de porcentagem between the ZIP/PHDf como representado pela SEQ ID N°: 233 and ZIP/PHDf do polipeptídeos PHDf-HD listadas na Tabela D1, and in Table D3 the identidade de porcentagem between the HD como representado pela SEQ ID N°: 234 and the HD do polipeptídeos PHDf-HD listadas na Tabela D1 do Exemplo 29.

Além disso, as sequências de aminoácido enriquecidas em aminoácidos básicos (Lys e Arg) ou ácidos (Glu e Asp), respectivamente denominados extensões básicas e ácidas, também podem ser facilmente identificadas simplesmente pela inspeção ocular (Figura 17). Alternativamente, a composição de aminoácido primária (em %) podem ser calculada spara determinar se um domínio de polipeptídeo é rico em aminoácidos específicos usando-se programas de software do servidor ExPASy, em particular, a ferramenta ProtParam (Gasteiger E *et al.* (2003) ExPASy: o servidor proteomics para o conhecimento de proteína em profundidadee análise. Nucleic Acids Res 31:3784-3788). A composição da

proteína de interesse pode ser então comparada com a composição de aminoácido média (in %) no banco de de Sequência de Proteína Swiss-Prot.

As espirais enroladas são importantes para a identificação de interações de proteína-proteína, tais como oligomerização, de proteínas idênticas, de proteínas da mesma família ou de proteínas não relacionadas. Um polipeptídeo PHDf-HD representa pelo menos uma região de espiral enrolada predita. Recentemente muito progresso foi feito na predição computacional de espirais enroladas dos dados de sequncia. Entre os algoritmos bem conhecidos por uma pessoa habilitada na técnica estão disponíveis nas ferramentas ExPASy Proteomics COILS, PAIRCOIL, PAIRCOIL2, MULTICOIL ou MARCOIL, hospedadas pelo Swiss Institute for Bioinformatics. No exemplo 33 e Figura 16, são conhecidos respectivamente os resultados numéricos e gráficos da SEQ ID N°: 180 como produzido pela análise de algoritmo COILS. Dois domínios de espiral enrolada preditos no terminal N são identificados em uma sequência de polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180, com uma probabilidade forte. Uma espiral enrolada de terminal C também é predita, com uma probalidade menor.

A tarefa da predição de localização subcelular de proteína é importante e bem estudada. O conhecimento de uma localização de proteína ajuda a elucidar a sua função. Os métodos experiemntais para a localização de proteína variam a partir da imunolocalização para a rotulação de proteínas usando-se proteína fluorescente verde (GFP). Tais métodos são precisos embora trabalho intensivo comparado com os métodos computacionais. Recentemente, muito progresso foi feito na predição computacional de localização de proteína a partir de dados de sequências. Entre os algoritmos bem conhecidos por uma pessoa habilitada na técnica estão disponíveis nas ferramentas ExPASy Proteomics hospedados pelo Swiss Institute for Bioinformatics, por exemplo, PSort, TargetP, ChloroP, LocTree, Predotar,

LipoP, MITOPROT, PATS, PTS1, SignalP e outros. A identificação da localização subcelular do polipeptídeo da invenção é mostrado no Exemplo 34. Em particular, a SEQ ID N°: 180 da presente invenção é designada ao compartimento nuclear de células eucarióticas.

Além disso, os polipeptídeos PHDf-HD úteis nos métodos da presente invenção (pelo menos em sua forma natural) tipicamente, mas não necessariamente, têm atividade reguladora transcripcional e capacidade para interagir com outras proteínas. Portanto, os polipeptídeos PHDf-HD com atividade reduzida, sem atividade reguladora transcripcional, com capacidade de interação proteína-proteína reduzida ou sem nenhuma capacidade de interação proteína-proteína, pode ser igualmente útil nos métodos da presente invenção. A atividade de ligação de DNA e as interações proteína-proteína podem ser facilmente determinadas in vitro ou in vivo usando-se técncias bem conhecidas na técnica (por exemplo in Current Protocols in Molecular Biology, Volumes 1 e 2, Ausubel et al. (1994), Current Protocols). Para determinar a atividade de ligação de DNA de polipeptídeos PHDf-HD, diversos ensaios são disponíveis, tais como ensaios de mudança de gel de ligação de DNA (ou ensaios de retardo de gel; Korfhage et al. (1994) Plant C 6: 695-708), ensaios de ligação de DNA in vitro (Schindler et al. (1993) Plant J 4(1): 137-150) ou ativação transcripcional de polipeptídeos PHDf-HD em levedura, células animais ou vegetais (Halbach et al. (2000) Nucleic Acid Res 28(18): 3542-3550). As sequências de ligação de DNA específicas podem ser determinadas usando-se a técnica de seleção de oligonucleotídeo aleatória (Viola & Gonzalez (May 26, 2007) Biochemistry).

A presente invenção é ilustrada transformando-se as plantas com a sequência de ácido nucleico representado pela SEQ ID N°: 179, que codifica a sequência de polipeptídeo da SEQ ID N°: 180. Entretanto, o desempenho da invenção não está restrito a estas sequências; os métodos da invenção podem ser vantajosamente realizados usando-se qualquer ácido

nucleico sequence que codificam PHDf-HD ou polipeptídeo PHDf-HD como definido neste.

Os exemplos de sequências de ácido nucleico que codificam polipeptídeos PHDf-HD são dados na Tabela D1 do Exemplo 29 neste. Tais sequências de ácido nucleico são úteis na realização dos métodos da invenção. As sequências de polipeptídeo dadas na Tabela D1 do Exemplo 29 são sequências de exemplos de ortólogos e parálogos do polipeptídeo PHDf-HD representado pela SEQ ID N°: 180, os termos "ortólogos" e "parálogos" sendo como definido neste. Ortólogos e parálogos adicionais podem ser facilmente definidos pela realização de uma denominada pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envovle submeter uma sequência em dúvida ao BLAST (por exemplo usando-se qualquer uma das sequências listadas na Tabela D1 do Exemplo 29) contra qualquer banco de dados de sequência, tal como os bancos de dados NCBI publicamente disponíveis. BLASTN ou TBLASTX (usando-se os valores default padrão) são, no geral, usados quando começa-se a partir de uma sequência de nucleotídeo e BLASTP ou TBLASTN (usando-se os valores default padrão) quando começa-se a partir de uma sequência de proteína.

Os resultados de BLAST podem ser opcionalmente filtrados. As sequências de comprimento total de cada um dos resultados filtrados ou resultados não filtrados são então submetidos ao BLAST novamente (segundo BLAST) contra sequências do organismo a partir do qual a sequência em questão é derivada (quando a sequência em questão é SEQ ID N°: 179 ou SEQ ID N°: 180, o segundo BLAST portanto, deve ser contra as sequências de arroz). Os resultados do primeiro e segundo BLASTs são então comparados. Um parálogo é identificado se um acerto de classificação alta do primeiro blast for da mesma espécie como a partir da qual a sequência em questão é derivada, submetendo-se ao BLAST novamente então resulta idealmente na sequência em questão entre os acertos mais altos; um ortólogo

é identificado se um acerto de classificação alta no primeiro BLAST não for da mesma espécie como a partir da qual a sequência em questão é derivada e preferivelmente resulta no BLAST novamente na sequência em questão estando entre os acertos mais altos.

Os acertos de classificação alta são aqueles tendo um valor E baixo. Quanto mais baixo o valor E, mais significante o registro (ou, em outras palavras, menor a chance que o acerto seja encontrado por chance). A computação do valor E é bem conhecido na técnica. Além dos valores E, comparações são registradas pela identidade de porcentagem. A identidade de porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácidos nucleicos comparadas (ou polipeptídeo) em um comprimento particular. No caso de famílias grandes, ClustalW pode ser usado, seguido por uma árvore de união próxima, para ajudar a visualizar o agrupamento de genes relacionados e para identificar ortólogos e parálogos. Qualquer agrupamento de sequência dentro do grupo que compreende SEQ ID N°: 180 (circulado na Figura 29) deve ser considerado estar dentro da definição já mencionada de um polipeptídeo PHDf-HD e deve ser considerado adequado para o uso nos métodos da invenção.

As variantes de ácido nucleico também podem ser úteis na prática dos métodos da invenção. Os exemplos de tais variantes incluem sequências de ácido nucleico que codificam homólogos e derivados de qualquer um de uma sequência de polipeptídeos dadas na Tabela D1 do Exemplo 29, os termos "homólogo" e "derivado" sendo como definido neste. Também úteis nos métodos da invenção are sequências de ácido nucleico que codificam homólogos e derivados de ortólogos ou parálogos de qualquer um de a sequência de polipeptídeos dadas na Tabela D1 do Exemplo 29. Os homólogos e derivados úteis nos métodos da presente invenção têm substancialmente a mesma atividade biológica e funcional como a proteína não modificada da qual estes são derivados.

As variantes de ácido nucleico adicionais úteis na prática dos métodos da invenção incluem porções de sequências de ácido nucleico que codificam polipeptídeos PHDf-HD, sequências de ácido nucleico que hibridizam-se a sequências de ácido nucleico que codificam polipeptídeos PHDf-HD, variantes de união de sequências de ácido nucleico que codificam polipeptídeos PHDf-HD, variantes alélicas de sequências de ácido nucleico que codificam polipeptídeos PHDf-HD e variantes de sequências de ácido nucleico que codificam polipeptídeos PHDf-HD obtidos pelo embaralhamento de gene. Os termos sequência de hibridização, variante de união, variante alélica e embaralhamento de gene são como descritos neste.

As sequências de ácido nucleico que codificam polipeptídeos PHDf-HD não necessitam ser sequências de comprimento total de ácido nucleico, visto que o desempenho dos métodos da invenção não contam com o uso de sequências de comprimento total de ácido nucleico. De acordo com a presente invenção, é fornecido um método para intensificar traços relacionados com o rendimento, preferivelmente que intensificam traços de sementes relacionados com o rendimento, em plantas, que compreende introduzir e expressar em uma planta uma porção de qualquer uma das sequências de ácido nucleico dadas na Tabela D1 do Exemplo 29 ou uma porção de uma sequência de ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer de uma sequência de polipeptídeos dadas na Tabela D1 do Exemplo 29.

Uma porção de uma sequência de ácido nucleico pode ser preparada, por exemplo, realizando-se uma ou mais anulações em uma sequência de ácido nucleico. As porções podem ser usadas na forma isolada ou estas podem ser fundidas a outras sequências codificadoras (ou não codificadoras) a fim de, por exemplo, produzir uma proteína que combine diversas atividades. Quando fundido a outras sequências codificadoras, o polipeptídeo resultante produzido na tradução pode ser maior do que aquele

predito para a porção de proteína.

As porções úteis nos métodos da invenção, codificam um polipeptídeo PHDf-HD como definido neste têm substancialmente a mesma atividade biológica como sequências de polipeptídeo dadas na Tabela D1 do Exemplo 29. Preferivelmente, a porção é uma porção de qualquer uma das sequências de ácido nucleico dadas na Tabela D1 do Exemplo 29 ou é uma porção de uma sequência de ácido nucleico que codifica um ortólogo ou parálogo de qualquer um de a sequência de polipeptídeos dadas na Tabela D1 do Exemplo 29. Preferivelmente a porção está, na ordem crescente de preferência pelo menos 600, 800, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600, 2800 nucleotídeos de comprimento consecutivo, os nucleotídeos consecutivos sendo de qualquer uma das sequências de ácido nucleico dadas na Tabela D1 do Exemplo 29 ou de uma sequência de ácido nucleico que codifica um ortólogo ou parálogo de qualquer um de uma sequência de polipeptídeos dadas na Tabela D1 do Exemplo 29. Mais preferivelmente a porção é uma porção da a sequência de ácido nucleico de SEQ ID N°: 179. Preferivelmente, a porção codifica uma sequência de polipeptídeo que quando usada na construção de uma árvore filogenética HD, tal como o descrito na Figura 13, os agrupamentos com o grupo de polipeptídeos PHDf-HD que compreende a sequência de polipeptídeo representado pela SEQ ID N°: 180 em vez de qualquer outro grupo HD.

Uma outra variante de sequência de ácido nucleico útil nos métodos da invenção é uma sequência de ácido nucleico capaz de hidridizar, sob as condições de estringência reduzida, preferivelmente sob condições estringentes, com uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD como definido neste ou com uma porção como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas,

preferivelmente itensificando os traços de sementes relacionados com o rendimento, que compreende introduzir e expressar em uma sequência de ácido nucleico vegetal capaz de hibridizar a qualquer uma das sequências de ácido nucleico dadas na Tabela D1 do Exemplo 29 ou que compreende introduzir e expressar em uma sequência de ácido nucleico vegetal capaz de hibridizar a uma sequência de ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de ácido nucleico dadas na Tabela D1 do Exemplo 29.

As sequências de hibridização úteis nos métodos da invenção codificam um polipeptídeo PHDf-HD como definido nestee têm substancialmente a mesma atividade biológica como uma sequência de polipeptídeos dadas na Tabela D1 do Exemplo 29. Preferivelmente, a sequência de hibridização é capaz de hidridizar a qualquer uma das sequências de ácido nucleico dadas na Tabela D1 do Exemplo 29 ou a uma porção de qualquer uma destas sequências, uma porção sendo como definido acima ou em que a sequência de hibridização é capaz de hibridizar uma sequência de ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma sequência de polipeptídeos dadas na Tabela D1 do Exemplo 29. Mais preferivelmente, a sequência de hibridização é capaz de hibridizar a uma sequência de ácido nucleico como representado pela SEQ ID N°: 179 ou a uma porção deste.

Preferivelmente, a sequência de hibridização codifica uma sequência de polipeptídeo que quando usada na construção de uma árvore filogenética HD, tal como o descrito na Figura 13, agrupamentos com o grupo de polipeptídeos PHDf-HD que compreendem a sequência de polipeptídeo representado pela SEQ ID N°: 180 em vez de qualquer outro grupo HD.

Uma outra variante de sequência de ácido nucleico útil nos métodos da invenção é uma variante de união que codifica um polipeptídeo PHDf-HD como definido neste acima, a variante de união sendo como

definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, preferivelmente itensificando os traços de sementes relacionados com o rendimento, que compreende introduzir e expressar em uma planta de uma variante de união de qualquer uma das sequências de ácido nucleico dadas na Tabela D1 do Exemplo 29 ou a variante de união de uma sequência de ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma da sequência de polipeptídeos dadas na Tabela D1 do Exemplo 29.

As variantes de união preferidas são as variantes de união de uma sequência de ácido nucleico representado pela SEQ ID N°: 179 ou a variante de união de uma sequência de ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 180. Preferivelmente, a sequência de polipeptídeo codificada pela variante de união, quando usado na construção de uma árvore filogenética HD, tal como o descrito na Figura 13, agrupamentos com o grupo de polipeptídeos PHDf-HD que compreende a sequência de polipeptídeo representado pela SEQ ID N°: 180 em vez de qualquer outro grupo HD.

Uma outra variante de sequência de ácido nucleico útil na realização dos métodos da invenção é uma variante alélica de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD como definido neste acima, uma variante alélica sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, preferivelmente itensificando os traços de sementes relacionados com o rendimento, que compreende introduzir e expressar em uma planta de uma variante alélica de qualquer uma das sequências de ácido nucleico dadas na Tabela D1 do Exemplo 29 ou que compreende introduzir e expressar em uma planta de uma variante alélica de uma sequência de ácido nucleico que

codifica um ortólogo, parálogo ou homólogo de qualquer uma da sequência de polipeptídeos dadas na Tabela D1 do Exemplo 29.

As variantes alélicas úteis nos métodos da presente invenção tem substancialmente a mesma atividade biológica como o polipeptídeo da SEQ ID N°: 180 e qualquer uma da sequência de polipeptídeos descritos na Tabela D1 do Exemplo 29. As variantes alélicas existem na natureza, e abrange dentro dos métodos da presente invenção é o uso destes alelos naturais. Preferivelmente, a variante alélica é uma variante alélica da SEQ ID N°: 179 ou uma variante alélica de uma sequência de ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 180. Preferivelmente, a sequência de polipeptídeo codificada pela variante alélica, quando usado na construção de uma árvore filogenética HD, tal como o descrito na Figura 13, agrupa-se com os polipeptídeos PHDf-HD que compreende a sequência de polipeptídeo representado pela SEQ ID N°: 180 em vez de qualquer outro grupo HD.

A mistura do gene ou evolução direcionada também pode ser usada para gerar variantes das sequências de ácido nucleico codificadas por polipeptídeos PHDf-HD como definido acima; o termo "a mistura do gene" sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, preferivelmente itensificando os traços de sementes relacionados com o rendimento, que compreende introduzir e expressar em uma planta de uma variante de qualquer uma das sequências de ácido nucleico dadas na Tabela D1 do Exemplo 29 ou que compreende introduzir e expressar em uma planta de uma variante de uma sequência de ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de polipeptídeos dadas na Tabela D1 do Exemplo 29, que a sequência da variante do ácido nucleico é obtida pelo embaralhamento de gene.

Vantajosamente, a presente invenção fornece até agora ácido nucleico PHDf-HD não conhecido e sequências de polipeptídeos.

De acordo com uma outra forma de realização da presente invenção, é fornecido uma molécula de ácido nucleico isolada que compreende:

(i) um ácido nucleico representado pela SEQ ID N°: 242;

(ii) um ácido nucleico ou fragmento do mesmo que é complementar à SEQ ID N°: 242;

(iii) um ácido nucleico que codifica um polipeptídeo JMJC tendo, na ordem crescente de preferência, pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência à sequência de aminoácido dada na SEQ ID N°: 243;

(iv) um ácido nucleico capaz de hibridizar sob condições estringentes a qualquer um dos ácidos nucleicos dadas em (i), (ii) ou (iii) acima.

De acordo com uma outra forma de realização da presente invenção, portanto, é fornecido um polipeptídeo isolado que compreende:

(i) um ácido nucleico representado pela SEQ ID N°: 242;

(ii) um ácido nucleico ou fragmento do mesmo que é complementar à SEQ ID N°: 242;

(iii) um ácido nucleico que codifica um polipeptídeo JMJC tendo, na ordem crescente de preferência, pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência à sequência de aminoácido dada na SEQ ID N°: 243;

(iv) um ácido nucleico capaz de hibridizar sob condições estringentes a qualquer um dos ácidos nucleicos dadas em (i), (ii) ou (iii) acima.

Preferivelmente, a sequência de polipeptídeo codificada pela sequência variante do ácido nucleico obtido pelo embaralhamento de gene,

quando usado na construção de uma árvore filogenética HD, tal como o descrito na Figura 13, agrupamentos com o grupo de polipeptídeos PHDf-HD que compreende a sequência de polipeptídeo representado pela SEQ ID N°: 180 antes do que com qualquer outro grupo HD.

Além disso, as variantes da sequência de ácido nucleico também podem ser obtidas pela mutagênese direcionada ao local. Diversos métodos são disponíveis para atingir a mutagênese direcionada ao local, o mais comum sendo PCR com base em métodos (Current Protocols in Molecular Biology. Wiley Eds.).

As sequências de ácido nucleico codificadas por polipeptídeos PHDf-HD podem ser derivadas a partir de qualquer fonte artificial ou natural. A sequência de ácido nucleico pode ser modificada a partir da forma natural na composição e/ou ambiente genômico através da consideração da manipulação humana. Preferivelmente a sequência de ácido nucleico que codifica o polipeptídeo PHDf-HD é a partir de uma planta, ainda preferivelmente de uma planta monocotiledônea, mais preferivelmente da família Poaceae, mais preferivelmente a sequência de ácido nucleico é de *Oryza sativa.*

O desempenho dos métodos da invenção dá plantas tendo traços intensificados relacionados com o rendimento. Em particular, o desempenho dos métodos da invenção dá plantas tendo rendimento aumentado, especialmente rendimento aumentado de semente com relação às plantas de controle. Os termos "produção" e "rendimento de semente" são descritos em mais detalhes na seção de "definições" neste.

A referência neste aos traços intensificados relacionados com o rendimento é determinada entender um aumento na biomassa (peso) por uma ou mais partes de uma planta, que pode incluir partes acima do solo (coletáveis) e/ou partes abaixo do solo (coletáveis). Em particular, tais partes colhíveis são sementes, e o desempenho dos métodos da invenção resultam

em plantas tendo rendimento aumentado de semente relativo ao rendimento de semente de plantas de controle.

Tomando o milho como um exemplo, um aumento da produção pode ser manifestada como um ou mais do seguinte: aumento no número de plantas estabelecidas por hectare ou acre, um aumento no número de espigas por planta, um aumento no número de séries, número de grãos por série, peso dos grãos, peso dos grãos milhar, comprimento/diâmetro da espiga, aumento da taxa de enchimento da semente (que é o número de sementes enchidas dividida pelo número total de sementes e multiplicadas por 100), entre outros. Tomando o arroz como um exemplo, um aumento da produção pode manifestar por si mesmo como um aumento em um ou mais do seguinte: número de plantas por hectare ou acre, número de panículas por planta, número de espigas por panículas, número de flores (florzinhas) por panícula (que é expressado como uma razão do número de sementes enchidas no número de panícula primárias), aumento da taxa de enchimento da semente (que é o número de sementes enchidas dividida pelo número total de sementes e multiplicadas por 100), aumento no peso dos grãos milhar, entre outros.

A presente invenção fornece um método para aumentar a produção, especialmente rendimento de semente de plantas, com relação às plantas de controle, que o método compreende modular, preferivelmente aumentar, expressão em uma planta de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD como definido neste.

Visto que as plantas transgênicas de acordo com a presente invenção têm traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, é igualmente que estas plantas exibam uma taxa de desenvolvimento aumentado (durante pelo menos parte de seu ciclo de vida), relativo à taxa de desenvolvimento de plantas de controle em um estágio correspondente no seu ciclo de vida.

A taxa de desenvolvimento aumentada pode ser específica por uma ou mais partes de uma planta (incluindo sementes) ou pode estar em toda parte substancialmente de uma planta total. As plantas tendo uma taxa de desenvolvimento aumentada pode ter um ciclo de vida mais curto. O ciclo de vida de uma planta pode ser levado significar um período necessário para o desenvolvimento a partir da semente madura seca até o estágio onde a planta tem produzido as sementes maduras secas, similar ao material de partida. Este ciclo de vida pode ser influenciado pelos fatores tal como vigor precoce, taxa de desenvolvimento, índice de viço, período e velocidade de florescência do amadurecimento da semente. O aumento na taxa de desenvolvimento pode acontecer em um ou mais estágios no ciclo de vida de uma planta ou durante substancialmente ao ciclo de vida total da planta. A taxa aumentada de desenvolvimento durante os estágios precoces no ciclo de vida de uma planta pode refletir o vigor itensificado (precoce). O aumento na taxa de desenvolvimento pode alter o ciclo de coleta de uma planta permitindo as plantas serem semeadas depois e/ou mais cedo coletadas do que de outra maneira seria possível (um efeito similar pode ser obtido com o período de florescência precoce). Se a taxa de desenvolvimento é suficientemente aumentada, este pode deixar para a semeadura adicional de sementes das mesmas espécies de plantas (por exemplo semeadura e coletagem de plantas de arroz seguido pela semeadura e coletagem das plantas de arroz adicionais todas dentro de um período de desenvolvimento convencional). Similarmente, se a taxa de desenvolvimento é suficientemente aumentada, este pode deixar para a semeadura adicional de sementes de espécies de plantas diferentes (por exemplo a semeadura e coletagem de plantas de milho seguido pela, por exemplo, a semeadura e coletagem opcional de soja, batata ou qualquer outra planta adequada). Os períodos de coletagem adicional do mesmo rizoma no caso de algumas plantas de cultivo também pode ser possível. Alterando o ciclo de coleta de uma planta pode levar a um aumento na produção da

biomassa anual por acre (devido a um aumento no número de vezes (dito em um ano) que qualquer planta particular pode ser desenvolvida e coletada). Um aumento na taxa de desenvolvimento também pode permitir o cultivo de plantas transgênicas em uma ampla área geográfica do que suas contrapartes de tipo selvagem, visto as limitações territoriais para o desenvolvimento de uma lavoura são frequentemente determinados pelas condições ambientais adversas no período do plantio (temporada precoce) ou no período de coletagem (temporada tardia). Tais condições adversas podem ser evitadas se o ciclo de coleta seja mais curto. A taxa de desenvolvimento pode ser determinação pela derivação de vários parâmetros a partir de curvas de desenvolvimento, tais parâmetros podem ser: T-Mid (o período de retirada para as plantas atingirem 50 % do seu tamanho máximo) e T-90 (período de retirada para as plantas atingirem 90 % do seu tamanho máximo), entre outros. A taxa de desenvolvimento como definido neste não é levado significar à florescência precoce.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção dá plantas tendo uma taxa de desenvolvimento aumentado com relação às plantas de controle. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a taxa de desenvolvimento das plantas, que o método compreende modular, preferivelmente aumentar, expressão em uma planta de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD como definido neste.

Os traços itensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, ocorre se a planta está sob condições sem estresse ou se a planta está exposta a várias tensões comparadas ao desenvolvimento de plantas de controle sob condições comparáveis. As plantas tipicamente respondem à exposição à tensão pelo crescimento mais lentamente. Em condições de diversas tensões, a planta ainda pode interromper o desenvolvimento

completamente. A tensão branda de outra maneira é definida neste como sendo qualquer tensão que qual a planta é exposta que não resulta no interrompimento da planta a desenvolver-se completamente sem a capacidade de recuperar o desenvolvimento. A tensão branda no sentido da invenção leva a uma redução no desenvolvimento das plantas submetidas à tensão de menos do que 40 %, 35 % ou 30 %, preferivelmente menos do que 25 %, 20 % ou 15 %, mais preferivelmente menos do que 14 %, 13 %, 12 %, 11 % ou 10 % ou menos em comparação à planta de controle sob condições sem estresse. Devido as vantagens nas práticas da agricultura (irrigação, fertilização, tratamentos com pesticidas) diversas tensões não são frequentemente encontradas em plantas de cultivos cultivadas. Como uma consequência, o desenvolvimento compromissado induzido pela tensão branda é frequentemente uma característica indesejável para a agricultura. As tensões brandas são as tensões abióticas e/ou bióticas (ambientais) que a planta é exposta. As tensões abióticas podem ser devido à aridez ou excesso de água, tensão anaeróbica, tensão salina, toxicidade química, tensão oxidativa e temperaturas de frio, calor e congelamento. A tensão abiótica pode ser uma tensão osmótica causa pela tensão à água (particularmente devido à aridez), tensão salina, tensão oxidativa ou uma tensão iônica. As tensões bióticas são tipicamente aquelas tensões causadas pelos patogênios, tal como bactérias, vírus, fungos, nematóides e insetos. O termo condições de "não tensão" como usado neste são aquelas condições ambientais que permitem o ótimo desenvolvimento das plantas. As pessoas habilitadas na técnica são atentas das condições do solo normal e condições climáticas para uma localização dada.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições sem estresse ou sob condições de secura brandas tendo traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, com relação

às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas desenvolvidas sob condições sem estresse ou sob condições de secura brandas, que o método compreende modular, preferivelmente aumentar, expressão em uma planta de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD. As plantas com ótimas condições de desenvolvimento, cultivadas como condições sem estresse, tipicamente produzem na ordem crescente de preferência pelo menos 90 %, 87 %, 85 %, 83 %, 80 %, 77 % ou 75 % da produção média de tal planta em um dado ambiente. A produção média pode ser calculada na base da coleta e/ou temporada em uma localização dada. As pessoas habilitadas na técnica estão atentas das produções de rendimento médio de uma lavoura.

O desempenho dos métodos de acordo com a presente invenção resultam em plantas desenvolvidas sob condições de estresse abiótico tendo traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, com relação às plantas de controle desenvolvidas sob condições de tensão comparáveis. Como relatado em Wang et al. (Planta (2003) 218: 1-14), tensão abiótica leva a uma sére de mudanças moleculares, bioquímicas, fisiológicas e morfológicas que afetam adversamente o desenvolvimento da planta e produtividade. Aridez, salinidade, temperaturas extremas e tensão oxidativa são conhecidas serem interconectadas e podem induzir o desenvolvimento e dano celular através de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve particularmente um alto grau de "cruzamento" entre a tensão de aridez e alta tensão por salinidade. Por exemplo, a aridez e/ou salinização são manifestados principalmente como tensão osmótica, resultando no rompimento da homeostase e distribuição de íon na célula. A tensão oxidativa, que frequentemente acompanha temperatura

alta ou baixa, salinidade ou tensão de aridez, pode causar a desnaturação de proteínas estruturais e funcionais. Como uma consequência, estas diversas tensões ambientais frequentemente ativam os caminhos de sinalização celular similares e respostas celulares, tal como a produção de proteínas de tensão, super regulação de anti-oxidantes, acúmulo de solutos compatíveis e suspensão do desenvolvimento. Visto que diversas tensões ambientais ativam os caminhos similares, a exemplificação da presente invenção com a tensão de aridez não deve ser visto como uma limitação da tensão de aridez, mas maé como uma avaliação para indicar o envolvimento dos polipeptídeos PHDf-HD como definido acima, na itensificação de traços relacionados com o rendimento, preferivelmente itensificando os traços de sementes relacionados com o rendimento, com relação às plantas de controle desenvolvidas em condições de tensão comparáveis, nas tensões abióticas no geral.

O termo "tensão abiótica" como definido neste é determinada entender qualquer um ou mais de: tensão à água (devido à aridez ou excesso de água), tensão anaeróbica, tensão salina, tensão à temperatura (devido as temperaturas de calor, frio ou congelamento), tensão de toxicidade química e tensão oxidativa. De acordo com um aspecto da invenção, a tensão abiótica é uma tensão osmótica, selecionado de tensão à água, tensão salina, tensão oxidativa e tensão iônica. Preferivelmente, a tensão à água é tensão de aridez. O termo tensão salina não é restrito ao sal comum (NaCl), mas pode ser qualquer tensão causada por um ou mais de: NaCl, KCl, LiCl, MgCl2, MgCl2, entre outros.

O desempenho dos métodos da invenção dá plantas tendo traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, sob condições de estresse abiótico com relação às plantas de controle desenvolvidas em condições de tensão comparáveis. Portanto, de acordo com a presente

invenção, é fornecido um método para intensificar os traços relacionados com o rendimento, preferivelmente itensificando os traços de sementes relacionados com o rendimento, em plantas desenvolvidas sob condições de estresse abiótico, que o método compreende modular, preferivelmente aumentar, expressão em uma planta de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD. De acordo com um aspecto da invenção, a tensão abiótica é uma tensão osmótica, selecionado de um ou mais do seguinte: tensão à água, tensão salina, tensão oxidativa e tensão iônica.

Um outro exemplo de tensão ambiental abiótica é a disponibilidade reduzida de um ou mais nutrientes que necessitam ser assimilados pelas plantas para o crescimento e desenvolvimento. Por causa da forte influência da eficiência da utilização da nutrição na produção da planta e qualidade do produto, uma imensa quantidade de fertilizador é vertido em campos para otimizar o desenvolvimento o qualidade vegetal. A produtividade de plantas de origem comum é limitada pelos três nutrientes primários, fósforo, potássio e nitrogênio, que é usualmente o elemento de limitação da taxa no desenvolvimento da planta deste três. Portanto o maior elemento nutricional requerido para o desenvolvimento é nitrogênio (N). Este é um constituinte de composto importantes numerosos observados em células vivas, incluindo aminoácidos, proteínas (enzimas), ácidos nucleicos, e clorofila. 1,5 % a 2 % da origem da planta seca é nitrogênio e aproximadamente 16 % da proteína total da planta. Deste modo, a disponibilidade do nitrogênio é um maior fator limitante para o desenvolvimento e produção da planta de cultivo (Frink et al. (1999) Proc Natl Acad Sci USA 96(4): 1175-1180), e tem bem como um maior impacto no acúmulo da proteína e composição de aminoácido. Portanto, o maior interesse são as plantas de cultivo com traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, quando desenvolvidos sob condições de limitação de

nitrogênio.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições de deficiência do nutriente, particularmente sob condições de deficiência de nitrogênio, traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento, preferivelmente itensificando os traços de sementes relacionados com o rendimento, em plantas desenvolvidas sob condições de deficiência do nutriente, que o método compreende modular, preferivelmente aumentar, expressão em uma planta de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD. Deficiência do nutriente pode resultar a partir da perda ou excesso de nutrientes tal como nitrogênio, fosfatos e outros compostos contendo fósforo, potássio, cálcio, cadmío, magnésio, manganês, ferro e boro, entre outros.

A presente invenção abrange plantas ou partes destas (incluindo sementes) ou células destas obtidas pelos métodos de acordo com a presente invenção. As plantas ou partes destas ou células destas compreende um transgene de ácido nucleico que codifica um polipeptídeo PHDf-HD como definido acima.

A invenção também fornece as construções genéticas e vetores para facilitar a introdução e/ou expressão modulada, preferivelmente aumentada em plantas das sequências de ácido nucleico codificadas por polipeptídeos PHDf-HD. A construção de genes pode ser inserida nos vetores, que pode ser comercialmente disponível, adequado para a transformação em plantas e para a expressão do gene de interesse nas células transformadas. A invenção também fornece o uso de uma construção de gene como definido neste nos métodos da invenção.

Mais especificamente, a presente invenção fornece uma construção que compreende:

(a) uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD como definido acima;

(b) uma ou mais sequências de controle capazes de modular, preferivelmente aumentar, a expressão da sequência de ácido nucleico de (a) e opcionalmente

(c) uma sequência de terminação de transcrição.

Preferivelmente, a sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD é como definido acima. O termo "sequência de controle" e "sequência de terminação" são como definidos neste.

Preferivelmente, uma das sequências de controle de uma construção é um promotor constitutivo isolado a partir do genoma da planta. Um exemplo de um promotor constitutivo da planta é um promotor GOS2, preferivelmente um promotor GOS2 de arroz, mais preferivelmente um promotor GOS2 como representado pela SEQ ID N°: 235.

As plantas são transformadas com um vetor que compreende qualquer uma das sequências de ácido nucleico descritas acima. O técnico habilitado é bem atento aos elementos genéticos que devem estar presentes no vetor a fim de transformar sucessivamente, selecionar e propagar as células hospedeiras contendo a sequência de interesse. A sequência de interesse é operacionalmente ligada a uma ou mais sequências de controle (pelo menos a um promotor).

Vantajosamente, qualquer tipo de promotor, se natural ou sintético, pode ser usado para modular, preferivelmente aumentar a expressão da sequência de ácido nucleico. Um promotor constitutivo é particularmente útil nos métodos, preferivelmente um promotor constitutivo isolado a partir do genoma da planta. O promotor constitutivo de planta conduz a expressão de uma sequência codificadora em um nível que é em todos os exemplos

abaixo que obtido sob o controle de um promotor 35S CaMV.

Outros promotores específicos pelo órgão, por exemplo pela expressão preferida em folhas, hastes, tubérculos, meristemas, sementes (embrião e/ou endosperma), são úteis na realização dos métodos da invenção. Ver a seção "Definições" neste pelas definições de vários tipos de promotores.

Deve ser claro que a aplicabilidade da presente invenção não é restrito a uma sequência de ácido nucleico que codifica o polipeptídeo PHDf-HD representado pela SEQ ID N°: 179, ou é a aplicabilidade da invenção restrita à expressão de uma sequência de ácido nucleíco que codifica o polipeptídeo PHDf-HD quando conduzido por um promotor constitutivo.

Opcionalmente, uma ou mais sequências terminadoras podem ser usadas na construção introduzida em uma planta. Os elementos reguladores adicionais podem incluir os itensificadores de transcrição bem como tradução. Aqueles habilitados na técnica serão atentos das sequências itensificadoras e terminadoras que podem ser adequado para o uso na realização da invenção. Uma sequência de íntron também pode ser adicionado à região de não tradução de 5' (UTR) ou na sequência codificadora para aumentar a quantidade da mensagem madura que acumula no citosol, como descrito na seção das definições. Outra sequência de controles (perto do promotor, itensificador, silenciador, sequências de íntron, regiões 3'UTR e/ou 5'UTR) pode ser elementos de estabilização de RNA e/ou proteína. Tais sequências seria conhecida ou já pode ser obtida por uma pessoa habilitada na técnica.

As construções genéticas da invenção ainda podem incluir uma origem da sequência de replicação que é requerida pela manutenção e/ou reprodução em um tipo celular específico. Um exemplo é quando uma construção genética é requerida ser mantida em uma células bacteriana como um elemento genético episomal (por exemplo, moléculas de plasmídio e

cosmídio). As origens preferidas da reprodução incluem, mas não são limitadas a, o f1-ori e colE1.

Para a detecção da transferência sucessível das sequências de ácido nucleico como usado nos métodos da invenção e/ou seleção de plantas transgênicas que compreende estas sequências de ácido nucleico, é vantajoso para o uso dos genes marcadores (ou genes repórter). Portanto, a construção genética pode opcionalmente compreender um gene de marcador selecionável. Os marcadores selecionáveis são descritos em mais detalhes na seção "definições" neste.

É conhecido que na integração transiente ou estável das sequências de ácido nucleico em células de planta, apenas uma maioria das células absorvem o DNA estranho e, se desejado, integram neste genoma, dependendo do vetor de expressão usado na técnica de trasnfecção usada. Para identificar e selecionar estes integrantes, um gene codificador para um marcador selecionável (tal como um descrito acima) é usualmente introduzido nas células hospedeiras junto com o gene de interesse. Estes marcadores podem ser por exemplo usados em mutantes em que estes genes não são funcionais por, por exemplo, anulação pelos métodos convencionais. Além disso, as moléculas da sequência de ácido nucleico que codifica um marcador selecionável pode ser introduzida em uma célula hospedeira no mesmo vetor que compreende a sequência codificadora dos polipeptídeos da invenção ou usados nos métodos da invenção ou senão em um vetor separado. As células que foram estavelmente transfectadas com a sequência do ácido nucleico introduzida pode ser identificada por exemplo pela seleção (por exemplo, células que tem integrado o marcador selecionável que sobrevive considerando a outras células mortas). Os genes marcadores podem ser removidos ou taxados a partir da célula transgênica uma vez que estes não são necessariamente mais longos. As técnicas para a remoção do gene marcador são conhecidas na técnica, técnicas úteis são descritas acima na seção

definições.

A invenção também fornece um método para a produção de plantas transgênicas tendo traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, com relação às plantas de controle, que compreende introdução e a expressão em uma planta de qualquer sequência do ácido nucleico que codifica um polipeptídeo PHDf-HD como definido neste acima.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, com relação às plantas de controle, que o método compreende:

(i) introduzir e expressar em uma planta, parte de planta ou células de planta uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD, sob o controle de promotor constitutivo de planta; e

(ii) cultivar a célula de planta sob condições que promovam o crescimento e desenvolvimento de planta.

A sequência de ácido nucleico de (i) pode ser qualquer uma das sequências de ácido nucleico capaz de codificar um polipeptídeo PHDf-HD como definido neste.

A sequência de ácido nucleico pode ser introduzido diretamente em uma célula de planta ou na própria planta (incluindo a introdução em um tecido, órgão ou qualquer outra parte de uma planta). De acordo com uma característica preferida da presente invenção, a sequência de ácido nucleico é preferivelmente introduzido em uma planta pela transformação. O termo "transformação" é descrito em mais detalhes na seção "definições" neste.

As células de planta geneticamente modificadas podem ser regenerados por intermédio de todos os métodos com que o trabalhador

habilitado seja familiar. Os métodos adequados podem ser observados nas publicações anteriomente mencionadas por S.D. Kung and R. Wu, Potrykus or Hofgen and Willmitzer.

No geral após a transformação, as células de planta ou agrupamentos celulares são selecionados pela presença de um ou mais marcadores que são codificados pelos genes expressíveis das plantas co-transferidos com o gene de interesse, seguindo que o material transformado é regenerado em um total da planta. Para selecionar as plantas transformadas, o material vegetal obtido na transformação é, como uma regra, submetido às condições seletivas de modo que as plantas transformadas podem ser distinguidas das plantas transformadas. Por exemplo, as sementes obtidas na maneira descrita acima pode ser plantada e, após um período de crescimento inicial, submetido a uma seleção adequada pela pulverização. Uma possibilidade adicional consiste no crescimento das sementes, se apropriado após a esterilização, em placas ágar usando um agente de seleção adequado de modo que apenas as sementes transformadas podem desenvolver-se nas plantas.

Alternativamente, as plantas transformadas são avaliadas pela presença de um marcador selecionável tal como um descrito acima.

Seguindo a transferência e regeneração de DNA, as plantas putativamente transformadas também podem ser avaliadas, por exemplo usando a análise Southern, pela presença do gene de interesse, número de cópia e/ou organização genômica. Alternativamente ou adicionalmente, os níveis de expressão do DNA introduzido novamente pode ser monitorado usando a análise Northern e/ou Western, ambas as técnicas sendo bem conhecidas às pessoas tendo habilidades comuns na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de meios, tal como pela propagação clonal ou técnicas de geração clássica. Por exemplo, uma primeira geração de planta transformada

(ou T1) pode ser por si mesmo e transformantes da segunda geração de homozigotos (ou T2) selecionados, e as plantas T2 ainda podem ser então propagado através de técnicas de geração clássicas. Os organismos transformados gerados podem tornar-se uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e células não transformadas; transformantes clonais (por exemplo, todas as células transformadas por conter o cassete de expressão); enxertos de tecidos transformados e não transformados (por exemplo, em plantas, um rizoma transformado enxertado em uma muda não transformada).

A presente invenção claramente estende-se a qualquer célula da planta ou planta produzida por qualquer um dos métodos descritos neste, e por todas partes da planta e propagules destes. A presente invenção estende-se ainda por abranger a progênie de uma célula transfectada ou transformada primária, tecido, órgão ou planta total que foi produzida por qualquer um dos métodos anteriomente mencionados, apenas o requerimento sendo que a progênie exibe as mesmas características fenotípicas e/ou genotípicas como aquelas produzidas pelo padrão nos métodos de acordo com a invenção.

A invenção também inclui as células hospedeiras contendo uma sequência do ácido nucleico isolada que codifica um polipeptídeo PHDf-HD como definido neste acima, operavelmente ligado a um promotor constitutivo de planta. As células hospedeiras preferidas de acordo com a invenção são células de planta. As plantas hospedeiras pelas sequências de ácido nucleico ou o vetor usado no método de acordo com a invenção, o cassete de expressão ou construção ou vetor são, à princípio, vantajosamente todas as plantas, que são capazes de sintetizar os polipeptídeos usados no método inventivo.

Os métodos da invenção são vantajosamente aplicáveis por qualquer planta. As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas, que pertencem à super família

Viridiplantae, em particular, plantas monocotiledônias e dicotiledônias incluindo forragem ou legumes de forragem, plantas ornamentais, lavouras de alimento, árvores e arbustos. De acordo com uma forma de realização preferida da presente invenção, a planta é uma planta de cultivo. Exemplos de plantas de cultivo incluem soja, girassol, canola, alfafa, colza, algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta monocotiledônia. Exemplos de plantas monocotiledônias incluem cana de açúcar. Mais preferivelmente a planta é um cereal. Exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

A invenção também estende-se as partes colhíveis de uma planta que compreendem uma sequência do ácido nucleico isolada que codifica um PHDf-HD (como definido neste acima) operacionalmente ligado a um promotor constitutivo de planta, tal como, mas não limitado a sementes, folhas, frutos, flores, hastes, rizomas, tubérculos e bulbos. A invenção além disso relata os produtos derivados, preferivelmente diretamente derivados, a partir da parte coletável de uma tal planta, tal como grânulos e pós secos, óleo, gordura e ácidos graxos, amido ou proteínas.

De acordo com uma característica preferida da invenção, a expressão modulada é a expressão aumentada. Os métodos para aumentar a expressão das sequências de ácido nucleico ou genes ou produtos dos genes, são bem documentados na técnica e exemplos são fornecidos na seção de definições.

Como mencionado acima, um método preferido para modular, preferivelmente aumentar, a expressão de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD é pela introdução e expressão em uma planta de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD; entretanto os efeitos da realização do método, isto é itensificando os traços relacionados com o rendimento, também podem ser atingido usando outras técnicas bem conhecidas, incluindo mas não limitado a rotulação de

ativação de T-DNA, TILLING, recombinação homóloga. Uma descrição destas técnicas são fornecidas na seção de definições.

A presente invenção também abrange o uso das sequências de ácido nucleico codificadas por polipeptídeos PHDf-HD como descrito neste e uso destes polipeptídeos PHDf-HD na itensificação de qualquer um dos traços mencionados acima relacionados com o rendimento, preferivelmente traços de sementes relacionados com o rendimento, em plantas.

As sequências de ácido nucleico que codifica o polipeptídeo PHDf-HD descrito neste ou os polipeptídeos PHDf-HD por si mesmos, pode observar nos programas de geração em que um marcador de DNA é identificado que pode ser geneticamente ligado a um gene que codifica o polipeptídeo PHDf-HD. Os genes/ sequências de ácido nucleico ou os polipeptídeos PHDf-HD por si mesmos podem ser usados para definir um marcador molecular. Este marcador de DNA ou proteína então podem ser usados nos programas da geração para selecionar as plantas tendo traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, como definido neste acima nos métodos da invenção.

As variantes alélicas de um gene/sequência do ácido nucleico que codifica um polipeptídeo PHDf-HD também podem ser observado em programas de geração assistidos pelo marcador. Tais programas de geração algumas vezes requerem a introdução da variação alélica pelo tratamento mutagênico de plantas, usando por exemplo a mutagênese EMS; alternativamente, o programa pode comerçar com uma coleção de variantes alélicas denominadas de origem "natural" causadas não itensionalmente. A identificação das variantes alélicas então acontece, por exemplo, por PCR. Este é seguido por uma etapa para a seleção de variantes alélicas superiores da sequência em questão e que dá traços intensificados relacionados com o rendimento. A seleção é tipicamente realizada pelo desempenho do

desenvolimento do monitoramento de plantas contendo variantes alélicas diferentes da sequência em questão. O desempenho do desenvolvimento pode ser monitorado em uma estufa ou no campo. As etapas opcionais adicionais incluem cruzamento de plantas em que a variante alélica superior foi indetificada com uma outra planta. Esta deve ser usada, por exemplo, para fabricar uma combinação de características fenotípicas interessantes.

As sequências de ácido nucleico codificadas por polipeptídeos PHDf-HD também podem ser usadas como sondas para o mapeamento físico e genético dos genes que estes estão a parte de, e como marcadores para traços ligados aqueles genes. Tal informação pode ser útil na geração de planta a fim de desenvolver as linhas com fenotipos desejados. Tal uso das sequências de ácido nucleico que codifica um polipeptídeo PHDf-HD requer apenas uma sequência de ácido nucleico de pelo menos 15 nucleotídeos de comprimento. As sequências de ácido nucleico que codifica um polipeptídeo PHDf-HD pode ser usado como marcadores de polimorfismo de comprimento do fragmento de restrição (RFLP). Southern blots (Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico da planta digerido por restrição pode ser sondado com as sequências de ácido nucleico que codifica um polipeptídeo PHDf-HD. Os padrões de grupos resultantes podem então ser submetidos à análise genética usando programas de computador tal como MapMaker (Lander et al. (1987) Genomics 1: 174-181) a fim de construir um mapa genético. Além disso, as sequências de ácido nucleico podem ser usadas para sondar Southern blots contendo DNA genômico tratado por endonuclease por restrição de uma série de indivíduos que representam o padrão e progênie de um cruzamento genético definido. A segregação do polimorfismo de DNA é notado e usado para calcular a posição de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD no mapa genético previamente obtido usando esta população (Botstein *et al.* (1980) Am. J. Hum. Genet. 32: 314-331).

A produção e uso de sondas derivadas do gene vegetal para o uso no mapeamento genético é descrito em Bernatzky and Tanksley (1986) Plant Mol. Biol. Reporter 4: 37-41. Numerosas publicações descrevem o mapeamento genético dos clones de cDNA específicos usando a metodologia resumida acima ou variações destas. Por exemplo, as populações intercruzadas F2, populações cruzadas novamente, populações aleatoriamente auxiliares, linhas isogênicas próximas, e outras séries de indivíduos podem ser usados para o mapeamento. Tais metodologias são bem conhecidas àqueles habilitados na técnica.

As sondas da sequência de ácido nucleico também podem ser usadas para o mapeamento físico (isto é, disposição das sequências nos mapas físicos; ver Hoheisel et al. In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346, e referências citadas neste).

Em uma outra forma de realização, as sondas da sequência de ácido nucleico podem ser usadas na fluorescência direta no mapeamento da hibridização in situ (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora os métodos correntes do mapeamento FISH favorece o uso de amplos clones (diversos kb à diversas centenas de kb; ver Laan et al. (1995) Genome Res. 5:13-20), melhoramento na sensibilidade pode permitir o desempenho do mapeamento FISH usando sondas mais curtas.

Uma variedade da amplificação da sequência do ácido nucleico com base em métodos para o mapeamento físico e genético pode ser realizado usando as sequências de ácido nucleico. Exemplos incluem amplificação específica por alelos (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismo dos fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332), ligação específica por alelos (Landegren et al. (1988) Science 241:1077-1080), reações de extensão por nucleotídeos (Sokolov (1990) Nucleic acid sequence Res. 18:3671), mapeamento híbrido

de radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e mapeamento Happy (Dear and Cook (1989) Nucleic acid sequence Res. 17:6795-6807). Para estes métodos, a sequência de uma sequência de ácido nucleico é usada para projetar e produzir pares de iniciadores para o uso na reação de amplificação ou nas reações da extensão iniciadora. O projeto de tais iniciadores é bem conhecido àqueles habilitados na técnica. Em métodos que utilizam mapeamento genético com base em PCR, este pode ser necessário para identificar as diferenças da sequência de DNA entre os padrões do cruzamento do mapeamento na região correspondente à sequência do ácido nucleico imediata. Este, entretanto, não é necessário no geral para os métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo traços intensificados relacionados com o rendimento, preferivelmente traços de semente intensificados relacionados com o rendimento, como descrito anteriormente acima. Estes traços também podem ser combinados com outros traços economicamente vantajosos, tal como traços que itensificam a produção adicional, tolerância a outras tensões bióticas e abióticas, traços que modificam várias características arquiteturais e/ou características bioquímicas e/ou fisiológicas.

V. Semelhante à bHLH11 (hélice 11-arco-hélice básica)

Surpreendentemente, foi observado agora que modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 dá plantas tendo traços intensificados relacionados com o rendimento com relação às plantas de controle. De acordo com uma primeira forma de realização, a presente invenção fornece um método para itensificar os traços relacionados com o rendimento em plantas com relação às plantas de controle, que compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11.

Um método preferido para modular (preferivelmente,

aumentar) a expressão de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 é pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11.

Qualquer referência à seguir a uma "proteína útil nos métodos da invenção" é determinada entender um polipeptídeo semelhante à bHLH11 como definido neste. Qualquer referência à seguir a um "ácido nucleico útil nos métodos da invenção" é determinada entender um ácido nucleico capaz de codificar um tal polipeptídeo semelhante à bHLH11. O ácido nucleico a ser introduzido em uma planta (e, portanto, útil na realização dos métodos da invenção) é qualquer ácido nucleico que codifica o tipo de proteína que será descrito agora, à seguir também denominado "ácido nucleico semelhante à bHLH11" ou "gene semelhante à bHLH11".

Um "polipeptídeo semelhante à bHLH11" como definido neste refere-se a qualquer polipeptídeo que compreende um domínio básico seguido por um domínio HLH (HMMPFam PF00010, ProfileScan PS50888, SMART SM00353) neste formando um domínio Hélice-Laço-Hélice básico (bHLH) (Interpro IPR001092). Preferivelmente, o polipeptídeo semelhante à bHLH11 compreende pelo menos um, preferivelmente dois, mais preferivelmente três, mais preferivelmente quatro ou mais dos seguintes motivos:

Motivo 1 (SEQ ID N°: 246):
(E/D)(D/S/E)(F/M)(L/F)(D/E/Q/L)(Q/H/E)

Motivo 2 (SEQ ID N°: 247):
RA(R/I/Q)RG(Q/H)ATDPHSIAER

Motivo 3 (SEQ ID N°: 248):
(M/I/V/L)(K/R)(A/S/Q/D/N)LQ(E/D/V)LVP

Motivo 4 (SEQ ID N°: 249):
(M/I)(L/I)DEI(I/V/L)(D/E/G)Y(V/L/I)(K/R)FL(Q/R)LQ(V/I)K

Motivo 5 (SEQ ID N°: 250): (V/I)LSMSR(L/V)G

Motivo 6 (SEQ ID N°: 251):

V(A/V/L/I)(K/R)(L/M)(M/L)(E/D)(E/D/S/K/T)(D/N/S)(M/V/I)(G/T/I)XAMQ(Y/L/F)L

em que X podem ser qualquer aminoácido, mas, preferivelmente um de S, T, A, M, K, N

Motivo 7 (SEQ ID N°: 252): (M/V)(P/S)(I/V)(S/T/A)LA

Alternativamente, o homólogo de uma proteína semelhante bHLH11 tem, na ordem crescente de preferência pelo menos 25 %, 26 %, 27 %, 28 %, 29 %, 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % ou 99 % da identidade da sequência total ao aminoácido representado pela SEQ ID N°: 245, fornecido que a proteína homóloga compreende os motivos conservados como resumido acima. A identidade da sequência total é determinada usando um algorítmo de alinhamento global, tal como o algorítmo Needleman Wunsch no programa GAP (GCG Wisconsin Package, Accelrys), preferivelmente com parâmetros default. Em comparação com a identidade da sequência total, a identidade de sequência no geral será mais alta quando apenas os domínios conservados ou motivos são considerados. A conservação da sequência é muito maior na região do domínio bHLH (ver Tabela E3 no Exemplo 45 e Figura 21). Portanto o domínio bHLH é um bom critério para a definição do grupo das proteínas semelhantes bHLH11. Preferivelmente, o polipeptídeo semelhante à bHLH11 compreende as sequências do motivo 8 (SEQ ID N°: 253):

SIAERLRRERIAERMRALQELVPNTNKTDRAVMLDEILD YVKFLRLQVKVL,

ou uma sequência que tem, na ordem crescente de preferência, pelo menos 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % ou 99 % de identidade de sequência à SEQ ID N°: 253. O domínio HLH como determinado por SMART estende-se sobre os resíduos 132 a 181 na SEQ ID N°: 245 e é compreendido no motivo 8.

Preferivelmente, a sequência de polipeptídeo que quando usado na construção de uma árvore filogenética, tal como o descrito na Figura 22, os grupos dentro do grupo das proteínas semelhantes bHLH11, antes do que com outras proteínas bHLH. Similarmente, as proteínas semelhante à bHLH11 de escolha agrupará dentro do subgrupo C quando uma árvore é construída de acordo com Figura 6 em Li et al. (2006), em vez de qualquer outro grupo.

Os termos "domínio", "sinal" e "motivo" são definidos na seção de "definições" neste. Os bancos de dados especializados existem para a identificação de domínios, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-244), InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318), Prosite (Bucher and Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004)) ou Pfam (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002)). Uma série de ferramentas para a análise *in silico* de sequências de proteínas está disponível no servidor proteômico ExPASy (Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31:3784-3788(2003)). Os domínios ou motivos também

podem ser identificados usando técnicas de rotina, tal como pelo alinhamento da sequência.

Os métodos para o alinhamento de sequências para comparação são bem conhecidos na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. GAP usa o algoritmo de Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar o alinhamento global (isto é medindo-se as sequências completas) de duas sequências que maximizam o número de equiparações e minimiza o número de fendas. O algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-10) calcula a identidade de sequência percentual e realiza uma análise estatística da similaridade entre as duas sequências. O software para realizar análise BLAST está publicamente disponível através do National Centre for Biotechnology Information (NCBI). Os homólogos podem ser facilmente identificados usando-se, por exemplo, o algoritmo de alinhamento de sequência múltipla ClustalW (versão 1.83), com os parâmetros de alinhamento default e um método de registro em porcentagem. As porcentagens globais de similaridade e de identidade também podem ser determinadas usando-se um dos métodos disponíveis no pacote de software MatGAT (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando-se proteína ou sequências de DNA.). A edição manual menor pode ser realizada para otimizar o alinhamento entre os motivos conservados, como deve estar evidente para uma pessa habilitada na técnica. Além disso, em vez de usar as sequências de comprimento total para a identificação de homólogos, domínios específicos também podem ser usados. Os valores de identidade de sequência podem ser determinados sobre o aminoácido inteiro ou sequência de aminoácido ou em domínios selecionados ou motivos conservados, usando-se os programas mencionados acima usando-se os parâmetros default. Para os alinhamentos locais, o algorítmo Smith-Waterman

é particularmente útil (Smith TF, Waterman MS (1981) J. Mol. Biol 147(1);195-7).

Além disso, os polipeptídeos semelhantes à bHLH11 (pelo menos em sua forma natural) tipicamente têm atividade de ligação de DNA. As ferramentas e técnicas para medir a atividade de ligação de DNA são bem conhecidas na técnica. Além disso, como mostrado na presente invenção, uma proteína semelhante à bHLH11, tal como a SEQ ID N°: 245, quando super expressadas em arroz, dá plantas tendo traços intensificados relacionados com o rendimento, em particular, taxa de enchimento aumentado. Detalhes adicionais são fornecidos na seção de Exemplos.

A presente invenção é ilustrada transformando-se as plantas com a sequência de ácido nucleico representado pela SEQ ID N°: 244, que codifica a sequência de polipeptídeo da SEQ ID N°: 245. Entretanto, o desempenho da invenção não está restrito a estas sequências; os métodos da invenção podem ser vantajosamente realizados usando-se qualquer ácido nucleico semelhante à bHLH11 ou polipeptídeo semelhante à bHLH11 como definido neste.

Exemplos de ácidos nucleicos que codificam o polipeptídeo semelhante à bHLH11 são dados na Tabela E1 do Exemplo 43 neste. Tais ácidos nucleicos são úteis na realização dos métodos da invenção. As sequências de aminoácido dadas na Tabela E1 do Exemplo 43 são sequências de exemplos de ortólogos e parálogos do polipeptídeo semelhante à bHLH11 representado pela SEQ ID N°: 245, os termos "ortólogos" e "parálogos" sendo como definido neste. Ainda os ortólogos e parálogos já podem ser identificados pela realização de uma denominada pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envolve o submetro de uma sequência em dúvida ao BLAST (por exemplo usando-se qualquer uma das sequências listadas na Tabela E1 do Exemplo 43) contra qualquer banco de dados de sequência, tal como os bancos de dados NCBI publicamente

disponíveis. BLASTN ou TBLASTX (usando-se os valores default padrão) são, no geral, usados quando começa-se a partir de uma sequência de nucleotídeo, e BLASTP ou TBLASTN (usando-se os valores default padrão) quando começa-se a partir de uma sequência de proteína. Os resultados de BLAST podem ser opcionalmente filtrados. As sequências de comprimento total de cada um dos resultados filtrados ou resultados não filtrados são então submetidos ao BLAST novamente (segundo BLAST) contra as sequências do organismo a partir do qual a sequência em questão é derivada (quando a sequência em questão é SEQ ID N°: 244 ou SEQ ID N°: 245, o segundo BLAST portanto, deve ser contra as sequências *Triticum aestivum*). Os resultados do primeiro e segundo BLAST são então comparados. Um parálogo é identificado se um acerto de classificação alta do primeiro blast da mesma espécie como a partir da qual a sequência em questão é derivada, submetendo-se ao BLAST novamente então resulta idealmente na sequência em questão entre os acertos mais altos; um ortólogo é identificado se um acerto de classificação alta no primeiro BLAST não da mesma espécie como a partir da qual a sequência em questão é derivada e preferivelmente resulta no BLAST novamente na sequência em questão estando entre os acertos mais altos.

Os acertos de classificação alta são aqueles tendo um valor E baixo. Quanto mais baixo o valor E, mais significante o registro (ou, em outras palavras, menor a chance que o acerto seja encontrado por chance). A computação do valor E é bem conhecido na técnica. Além dos valores E, as comparações são registradas pela identidade de porcentagem. A identidade de porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. No caso de famílias grandes, ClustalW pode ser usado, seguido por uma árvore de união próxima, para ajudar a visualizar o agrupamento de genes relacionados e para identificar ortólogos e parálogos.

As variantes de ácido nucleico também podem ser úteis na prática dos métodos da invenção. Exemplos de tais variantes incluem ácidos nucleicos que codificam os homólogos e derivados de qualquer uma das sequências de aminoácido dadas na Tabela E1 do Exemplo 43, os termos "homólogo" e "derivado" sendo como definido neste. Também são úteis nos métodos da invenção os ácidos nucleicos que codificam os homólogos e derivados de ortólogos ou parálogos de qualquer uma das sequências de aminoácido dadas na Tabela E1 do Exemplo 43. Os homólogos e derivados úteis nos métodos da presente invenção têm substancialmente a mesma atividade biológica e funcional como a proteína não modificada da qual estes são derivados.

As variantes de ácido nucleico adicionais úteis na prática dos métodos da invenção incluem as porções de ácidos nucleicos que codificam os polipeptídeos semelhantes à bHLH11, ácidos nucleicos que hibridizam-se ácidos nucleicos que codificam os polipeptídeos semelhantes à bHLH11, variantes de união de ácidos nucleicos que codificam os polipeptídeos semelhantes à bHLH11, variantes alélicas de ácidos nucleicos que codificam os polipeptídeos semelhantes à bHLH11 e variantes de ácidos nucleicos que codificam os polipeptídeos semelhantes à bHLH11 obtidos pelo embaralhamento de gene. Os termos sequência de hibridização, variante de união, variante alélica e embaralhamento de gene são como descritos neste.

Os ácidos nucleícos que codificam os polipeptídeos semelhantes à bHLH11 não necessitam ser os ácidos nucleicos de comprimento total, visto que o desempenho dos métodos da invenção não contam com o uso de sequências de comprimento total de ácido nucleico. De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta uma porção de qualquer uma das sequências de ácido nucleico dadas na Tabela E1 do Exemplo 43 ou uma porção de um

ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de aminoácido dadas na Tabela E1 do Exemplo 43.

Uma porção de um ácido nucleico pode ser preparada, por exemplo, realizando-se uma ou mais anulações no ácido nucleico. As porções podem ser usadas na forma isolada ou estas podem ser fundidas a outras sequências codificadoras (ou não codificadoras) a fim de, por exemplo, produzir uma proteína que combine diversas atividades. Quando fundido a outras sequências codificadoras, o polipeptídeo resultante produzido na tradução pode ser maior do que aquele predito para a porção de proteína.

As porções úteis nos métodos da invenção, codificam um polipeptídeo semelhante à bHLH11 como definido nestee têm substancialmente a mesma atividade biológica como as sequências de aminoácido dadas na Tabela E1 do Exemplo 43. Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos dados na Tabela E1 do Exemplo 43 ou é uma porção de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das sequências de aminoácido dadas na Tabela E1 do Exemplo 43. Preferivelmente a porção é pelo menos 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500 nucleotídeos consecutivos de comprimento, os nucleotídeos consecutivos sendo de qualquer uma das sequências de ácido nucleico dadas na Tabela E1 do Exemplo 43 ou de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das sequências de aminoácido dadas na Tabela E1 do Exemplo 43. Mais preferivelmente a porção é uma porção do ácido nucleico da SEQ ID N°: 244. Preferivelmente, a porção codifica um fragmento de uma sequência de aminoácido que, quando usado na construção de uma árvore filogenética, tal como o descrito na Figura 22, os grupos dentro do grupo das proteínas semelhantes bHLH11, antes do que com outras proteínas bHLH. Similarmente, as proteínas semelhante à bHLH11 de escolha agrupará dentro

do subgrupo C quando uma árvore é construída de acordo com Figura 6 em Li et al. (2006), em vez de qualquer outro grupo.

Uma outra variante de ácido nucleico útil nos métodos da invenção é um ácido nucleico capaz de hibridizar, sob condições de estringência reduzida, preferivelmente sob condições estringentes, com um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 como definido neste ou com uma porção como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta, um ácido nucleico capaz de hibridizar a qualquer um dos ácidos nucleicos dadas na Tabela E1 do Exemplo 43 ou que compreende introduzir e expressar em uma planta, um ácido nucleico capaz de hibridizar a um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de ácido nucleico dadas na Tabela E1 do Exemplo 43.

As sequência de hibridização úteis nos métodos da invenção codificam um polipeptídeo semelhante à bHLH11 como definido neste, tendo substancialmente a mesma atividade biológica como as sequências de aminoácido dadas na Tabela E1 do Exemplo 43. Preferivelmente, a sequência de hibridização é capaz de hibridizar o complemento de qualquer um dos ácidos nucleicos dados na Tabela E1 do Exemplo 43 ou a uma porção de qualquer uma destas sequências, uma porção sendo como definido acima ou a sequência de hibridização é capaz de hibridizar o complemento de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das sequências de aminoácido dadas na Tabela E1 do Exemplo 43. Mais preferivelmente, a sequência de hibridização é capaz de hibridizar o complemento de um ácido nucleico como representado pela SEQ ID N°: 244 ou a uma porção deste.

Preferivelmente, a sequência de hibridização codifica um

polipeptídeo com uma sequência de aminoácido que, quando o comprimento total e a construção usada de uma árvore filogenética, tal como o descrito na Figura 22, os grupos dentro do grupo das proteínas semelhantes bHLH11, antes do que com outras proteínas bHLH. Similarmente, as proteínas semelhante à bHLH11 de escolha agrupará dentro do subgrupo C quando uma árvore é construída de acordo com Figura 6 em Li et al. (2006), em vez de qualquer outro grupo.

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante de união que codifica um polipeptídeo semelhante à bHLH11 como definido neste acima, a variante de união sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta de uma variante de união de qualquer uma das sequências de ácido nucleico dadas na Tabela E1 do Exemplo 43 ou a variante de união de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de aminoácido dadas na Tabela E1 do Exemplo 43.

As variantes de união preferidas são as variantes de união de um ácido nucleico representado pela SEQ ID N°: 244 ou a variante de união de um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 245. Preferivelmente, a sequência de aminoácido codificada pela variante de união, quando usado na construção de uma árvore filogenética, tal como o descrito na Figura 22, os grupos dentro do grupo das proteínas semelhantes bHLH11, antes do que com outras proteínas bHLH. Similarmente, as proteínas semelhante à bHLH11 de escolha agrupará dentro do subgrupo C quando uma árvore é construída de acordo com Figura 6 em Li et al. (2006), em vez de qualquer outro grupo.

Uma outra variante de ácido nucleico útil na realização dos

métodos da invenção é uma variante alélica de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 como definido neste acima, uma variante alélica sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta de uma variante alélica de qualquer um dos ácidos nucleicos dados na Tabela E1 do Exemplo 43 ou que compreende introduzir e expressar em uma planta de uma variante alélica de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de aminoácido dadas na Tabela E1 do Exemplo 43.

Os polipeptídeos codificados pelas variantes alélicas úteis nos métodos da presente invenção tem substancialmente a mesma atividade biológica como o polipeptídeo semelhante à bHLH11 da SEQ ID N°: 244 e qualquer um dos aminoácidos descritos na Tabela E1 do Exemplo 43. As variantes alélicas existem na natureza, e abrangem dentro dos métodos da presente invenção é o uso destes alelos naturais. Preferivelmente, a variante alélica é uma variante alélica da SEQ ID N°: 244 ou uma variante alélica de um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 245. Preferivelmente, a sequência de aminoácido codificada pela variante alélica, quando usado na construção de uma árvore filogenética, tal como o descrito na Figura 22, os grupos dentro do grupo das proteínas semelhantes bHLH11, antes do que com outras proteínas bHLH. Similarmente, as proteínas semelhante à bHLH11 de escolha agrupará dentro do subgrupo C quando uma árvore é construída de acordo com Figura 6 em Li et al. (2006), em vez de qualquer outro grupo.

A mistura do gene ou evolução direcionada também pode ser usada para gerar variantes dos ácidos nucleicos que codificam os polipeptídeos semelhantes à bHLH11 como definido acima; o termo "a

mistura do gene" sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta de uma variante de qualquer uma das sequências de ácido nucleico dadas na Tabela E1 do Exemplo 43 ou que compreende introduzir e expressar em uma planta de uma variante de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das sequências de aminoácido dadas na Tabela E1 do Exemplo 43, que a variante do ácido nucleico é obtido pelo embaralhamento de gene.

Preferivelmente, a sequência de aminoácido codificada pela variante ácido nucleico obtido pelo embaralhamento de gene, quando usado na construção de uma árvore filogenética, tal como o descrito na Figura 22, os grupos dentro do grupo das proteínas semelhantes bHLH11, antes do que com outras proteínas bHLH. Similarmente, as proteínas semelhante à bHLH11 de escolha agrupará dentro do subgrupo C quando uma árvore é construída de acordo com Figura 6 em Li et al. (2006), em vez de qualquer outro grupo.

Além disso, as variantes do ácido nucleico também podem ser obtidos pela mutagênese direcionada ao local. Diversos métodos são disponíveis para atingir a mutagênese direcionada ao local, o mais comum sendo PCR com base em métodos (Current Protocols in Molecular Biology. Wiley Eds.).

Os ácidos nucleicos que codificam os polipeptídeos semelhantes à bHLH11 podem ser derivados a partir de qualquer fonte artificial ou natural. O ácido nucleico pode ser modificado a partir da forma natural na composição e/ou ambiente genômico através da consideração manipulação humana. Preferivelmente o polipeptídeo semelhante à bHLH11 que codifica o ácido nucleico é a partir de uma planta, ainda preferivelmente de uma planta monocotiledônea, mais preferivelmente da família Poaceae,

mais preferivelmente o ácido nucleico é de *Triticum aestivum.*

O desempenho dos métodos da invenção dá plantas tendo traços intensificados relacionados com o rendimento. Em particular, o desempenho dos métodos da invenção dá plantas tendo rendimento aumentado, especialmente rendimento aumentado de semente com relação às plantas de controle. Os termos "produção" e "rendimento de semente" são descritos em mais detalhes na seção de "definições" neste.

A referência neste aos traços intensificados relacionados com o rendimento é determinada entender um aumento na biomassa (peso) por uma ou mais partes de uma planta, que pode incluir partes acima do solo (coletáveis) e/ou partes abaixo do solo (coletáveis). Em particular, tais partes colhíveis são sementes, e o desempenho dos métodos da invenção resultam em plantas tendo rendimento aumentado de semente relativo ao rendimento de semente de plantas de controle.

Tomando o milho como um exemplo, um aumento da produção pode ser manifestada como um ou mais do seguinte: aumento no número de plantas estabelecidas por metro quadrado, um aumento no número de espigas por planta, um aumento no número de séries, número de grãos por série, peso dos grãos, peso dos grãos milhar, comprimento/diâmetro da espiga, aumento da taxa de enchimento da semente (que é o número de sementes enchidas dividida pelo número total de sementes e multiplicadas por 100), entre outros. Tomando o arroz como um exemplo, um aumento da produção pode manifestar por si mesmo como um aumento em um ou mais do seguinte: número de plantas por metro quadrado, número de panículas por planta, número de espigas por panículas, número de flores (florzinhas) por panícula (que é expressado como uma razão do número de sementes enchidas no número de panícula primárias), aumento da taxa de enchimento da semente (que é o número de sementes enchidas dividida pelo número total de sementes e multiplicadas por 100), aumento no peso dos grãos milhar, entre outros.

A presente invenção fornece um método para aumentar a produção, especialmente rendimento de semente de plantas, com relação às plantas de controle, que o método compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 como definido neste.

Visto que as plantas transgênicas de acordo com a presente invenção tem um rendimento aumentado, é igualmente que estas plantas exibam uma taxa de desenvolvimento aumentado (durante pelo menos parte de seu ciclo de vida), relativo à taxa de desenvolvimento de plantas de controle em um estágio correspondente no seu ciclo de vida.

A taxa de desenvolvimento aumentada pode ser específica por uma ou mais partes de uma planta (incluindo sementes) ou pode estar em toda parte substancialmente de uma planta total. As plantas tendo uma taxa de desenvolvimento aumentado pode ter um ciclo de vida mais curto. O ciclo de vida de uma planta pode ser levado significar um período necessário para o desenvolvimento a partir da semente madura seca até o estágio onde a planta tem produzido as sementes maduras secas, similar ao material de partida. Este ciclo de vida pode ser influenciado pelos fatores tal como vigor precoce, taxa de desenvolvimento, índice de viço, período e velocidade de florescência do amadurecimento da semente. O aumento na taxa de desenvolvimento pode acontecer em um ou mais estágios no ciclo de vida de uma planta ou durante substancialmente ao ciclo de vida total da planta. A taxa aumentada de desenvolvimento durante os estágios precoces no ciclo de vida de uma planta pode refletir no vigor itensificado. O aumento na taxa de desenvolvimento pode alterar o ciclo de coleta de uma planta permitindo as plantas serem semeadas depois e/ou mais cedo coletadas do que de outra maneira seria possível (um efeito similar pode ser obtido com período de florescência precoce). Se a taxa de desenvolvimento é suficientemente aumentada, este pode deixar para a semeadura adicional de sementes das mesmas espécies de

plantas (por exemplo semeadura e coletagem de plantas de arroz seguido pela semeadura e coletagem das plantas de arroz adicionais todas dentro de um período de desenvolvimento convencional). Similarmente, se a taxa de desenvolvimento é suficientemente aumentada, este pode deixar para a semeadura adicional de sementes de espécies de plantas diferentes (por exemplo a semeadura e coletagem de plantas de milho seguido pela, por exemplo, a semeadura e coletagem opcional de soja, batata ou qualquer outra planta adequada). Os períodos de coletagem adicional do mesmo rizoma no caso de algumas plantas de cultivo também podem ser possível. Alterando o ciclo de coleta de uma planta pode levar a um aumento na produção da biomassa anual por metro quadrado (devido a um aumento no número de vezes (dito em um ano) que qualquer planta particular pode ser desenvolvida e coletada). Um aumento na taxa de desenvolvimento também pode permitir o cultivo de plantas transgênicas em uma ampla área geográfica do que suas contrapartes de tipo selvagem, visto que as limitações territoriais para o desenvolvimento de uma lavoura são frequentemente determinados pelas condições ambientais adversas no período do plantio (temporada precoce) ou no período de coletagem (temporada tardia). Tais condições adversas podem ser evitadas se o ciclo de coleta seja mais curto. A taxa de desenvolvimento pode ser determinação pela derivação de vários parâmetros a partir de curvas de desenvolvimento, tais parâmetros podem ser: T-Mid (o período de retirada para as plantas atingirem 50 % do seu tamanho máximo) e T-90 (período de retirada para as plantas atingirem 90 % do seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção dá plantas tendo uma taxa de desenvolvimento aumentado com relação às plantas de controle. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a taxa de desenvolvimento de plantas, que o método compreende modular a

expressão em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 como definido neste.

Um aumento na produção e/ou taxa de desenvolvimento ocorre se a planta está sob condições sem estresse ou se a planta está exposta a várias tensões em comparação à plantas de controle. As plantas tipicamente respondem à exposição à tensão pelo crescimento mais lentamente. Em condições de diversas tensões, a planta ainda pode interromper o desenvolvimento completamente. A tensão branda de outra maneira é definida neste como sendo qualquer tensão que qual a planta é exposta que não resulta no interrompimento da planta desenvolver-se completamente sem a capacidade recuperar o desenvolvimento. A tensão branda no sentido da invenção leva a uma redução no desenvolvimento das plantas submetidas à tensão de menos do que 40 %, 35 % ou 30 %, preferivelmente menos do que 25 %, 20 % ou 15 %, mais preferivelmente menos do que 14 %, 13 %, 12 %, 11 % ou 10 % ou menos em comparação à planta de controle sob as condições sem estresse. Devido as vantagens nas práticas da agricultura (irrigação, fertilização, tratamentos com pesticidas) diversas tensões não são frequentemente encontradas em plantas de cultivos cultivadas. Como uma consequência, o desenvolvimento compromissado induzido pela tensão branda é frequentemente uma característica indesejável para a agricultura. As tensões brandas são as tensões abióticas e/ou bióticas (ambientais) que a planta é exposta. As tensões abióticas podem ser devido à aridez ou excesso de água, tensão anaeróbica, tensão salina, toxicidade química, tensão oxidativa e temperaturas de frio, calor e congelamento. A tensão abiótica pode ser uma tensão osmótica causa pela tensão à água (particularmente devido à aridez), tensão salina, tensão oxidativa ou uma tensão iônica. As tensões bióticas são tipicamente aquelas tensões causadas pelos patogênios, tal como bactérias, vírus, fungos, nematóides e insetos.

Em particular, os métodos da presente invenção podem ser

realizados sob condições sem estresse ou sob condições de aridez branda para dar plantas tendo rendimento aumentado com relação às plantas de controle. Como relatado em Wang et al. (Planta (2003) 218: 1-14), tensão abiótica leva a uma sére de mudanças moleculares, bioquímicas, fisiológicas e morfológicas que adversamente afeta o desenvolvimento da planta e produtividade. A aridez, salinidade, temperaturas extremas e tensão oxidativa estão conhecidas serem interconectadas e podem induzir o desenvolvimento e dano celular através de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve particularmente um alto grau de "cruzamento" entre a tensão de aridez e alta tensão por salinidade. Por exemplo, aridez e/ou salinização são manifestados principalmente como tensão osmótica, resultando no rompimento da homeostase e distribuição de íon na célula. A tensão oxidativa, que frequentemente acompanha temperatura alta ou baixa, salinidade ou tensão de aridez, pode causar a desnaturação de proteínas estruturais e funcionais. Como uma consequência, estas diversas tensões ambientais frequentemente ativam os caminhos de sinalização celular similares e respostas celulares, tal como a produção de proteínas de tensão, super regulação de anti-oxidantes, acúmulo de solutos compatíveis e suspensão do desenvolvimento. O termo condições de "não tensão" como usado neste são aquelas condições ambientais que permitem o ótimo desenvolvimento das plantas. As pessoas habilitadas na técnica são atentas das condições do solo normal e condições climáticas para uma localização dada. As plantas com ótimas condições de desenvolvimento, (desenvolvimento sob condições sem estresse) tipicamente produzem na ordem crescente de preferência pelo menos 90 %, 87 %, 85 %, 83 %, 80 %, 77 % ou 75 % da produção média de tal planta em um dado ambiente. A produção média pode ser calculada na base da coleta e/ou temporada. As pessoas habilitadas na técnica estão atentas das produções de rendimento médio de uma lavoura.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições sem estresse ou sob condições de secura brandas rendimento aumentado com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a produção em plantas desenvolvidas sob condições sem estresse ou sob condições de secura brandas, que o método compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições de deficiência do nutriente, particularmente sob condições de deficiência de nitrogênio, rendimento aumentado com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a produção em plantas desenvolvidas sob condições de deficiência do nutriente, que o método compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11, fornecido que a deficiência do nutriente não é uma deficiência de fosfato. A deficiência do nutriente pode resultar a partir da perda de nutrientes tal como nitrogênio, fosfatos e outros compostos contendo fósforo, potássio, cálcio, cadmío, magnésio, manganês, ferro e boro, entre outros. Entretanto, o termo "deficiência do nutriente" como usado no contexto da presente invenção não abrange uma deficiência em fosfato.

A presente invenção abrange plantas ou partes destas (incluindo sementes) obtidas pelos métodos de acordo com a presente invenção. As plantas ou partes destas compreendem um transgene de ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 como definido acima.

A invenção também fornece construções genéticas e vetores para facilitar introdução e/ou expressão em plantas de ácidos nucleicos que

codificam os polipeptídeos semelhantes à bHLH11. A construção de genes podem ser inseridos nos vetores, que podem ser comercialmente disponíveis, adequados para a transformação em plantas e adequados para a expressão do gene de interesse nas células transformadas. A invenção também fornece o uso de uma construção de gene como definido neste nos métodos da invenção.

Mais especificamente, a presente invenção fornece uma construção que compreende:

(a) um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 como definido acima;

(b) uma ou mais sequências de controle capazes de conduzir a expressão da sequência de ácido nucleico de (a) e opcionalmente

(c) uma sequência de terminação de transcrição.

Preferivelmente, o ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 é como definido acima. O termo "sequência de controle" e "sequência de terminação" são como definidos neste.

As plantas são transformadas com um vetor que compreende qualquer um dos ácidos nucleicos descritos acima. O técnico habilitado é bem atento aos elementos genéticos que devem estar presentes no vetor a fim de transformar sucessivamente, selecionar e propagar células hospedeiras contendo a sequência de interesse. A sequência de interesse é operacionalmente ligada a uma ou mais sequências de controle (pelo menos a um promotor).

Vantajosamente, qualquer tipo de promotor, se natural ou sintético, pode ser usado para conduzir a expressão de uma sequência de ácido nucleico, mas, preferivelmente o promotor é de origem vegetal. Um promotor constitutivo é particularmente útil nos métodos. Preferivelmente o promotor constitutivo também é um promotor ubíquo de força média. Ver a

seção de "Definições" neste para as definições de vários tipos de promotores.

Deve ser claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico que codifica o polipeptídeo semelhante à bHLH11 representado pela SEQ ID N°: 1, ou é a aplicabilidade da invenção restrita à expressão de um ácido nucleico que codifica o polipeptídeo semelhante à bHLH11 quando conduzido por um promotor constitutivo.

O promotor constitutivo é preferivelmente de um promotor de força média, tal como um promotor GOS2, preferivelmente o promotor é um promotor GOS2 de arroz. Ainda preferivelmente o promotor constitutivo é representado por uma sequência de ácido nucleico substancialmente similar à SEQ ID N°: 256, mais preferivelmente o promotor constitutivo é como representado pela SEQ ID N°: 256. Ver a seção das "Definições" neste por exemplos adicionais de promotores constitutivos.

Opcionalmente, uma ou mais sequências terminadoras podem ser usadas na construção introduzida em uma planta. Preferivelmente, a construção compreende um cassete de expressão que compreende o promotor GOS2 substancialmente similar à SEQ ID N°: 256 e o ácido nucleico que codifica o polipeptídeo semelhante à bHLH11.

Os elementos reguladores adicionais podem incluir a transcrição bem como os itensificadores de tradução. Aqueles habilitados na técnica serão atentos das sequências itensificadoras e terminadoras que podem ser adequados para o uso na realização da invenção. Uma sequência de íntron também podem ser adicionado à região de não tradução de 5' (UTR) ou na sequência codificadora para aumentar a quantidade da mensagem madura que acumula no citosol, como descrito na seção de definições. Outra sequência de controle (perto do promotor, itensificador, silenciador, sequências de íntron, regiões 3'UTR e/ou 5'UTR) pode ser elementos de estabilização de RNA e/ou proteína. Tais sequências seria conhecida ou já pode ser obtida por uma pessoa habilitada na técnica.

As construções genéticas da invenção ainda podem incluir uma origem da sequência de replicação que é requerida pela manutenção e/ou reprodução em um tipo celular específico. Um exemplo é quando uma construção genética é requerida ser mantida em uma células bacteriana como um elemento genético episomal (por exemplo, moléculas de plasmídio e cosmídio). As origens preferidas da reprodução incluem, mas não são limitadas a, o f1-ori e colE1.

Para a detecção da transferência sucessível das sequências de ácido nucleico como usado nos métodos da invenção e/ou seleção de plantas transgênicas que compreende estes ácidos nucleicos, é vantajoso para o uso dos genes marcadores (ou genes repórter). Portanto, a construção genética pode opcionalmente compreender um gene de marcador selecionável. Os marcadores selecionáveis são descritos em mais detalhes na seção de "definições" neste. Os genes marcadores podem ser removidos ou taxados a partir da célula transgênica uma vez que estes não são necessariamente mais longos. As técnicas para a remoção do marcador são conhecidas na técnica, técnicas úteis são descritos acima na seção de definições.

A invenção também fornece um método para a produção de plantas transgênicas tendo traços intensificados relacionados com o rendimento com relação às plantas de controle, que compreende introdução e expressão em uma planta de qualquer ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 como definido neste acima.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo traços intensificados aumentados relacionados com o rendimento, particularmente produção aumentada (semente), que o método compreende:

(i) introduzir e expressar em uma planta ou célula de planta um ácido nucleico que codifica o polipeptídeo semelhante à bHLH11; e

(ii) cultivar a célula de planta sob condições que promovam o

crescimento e desenvolvimento de planta.

O ácido nucleico de (i) pode ser qualquer um dos ácidos nucleicos capazes de codificar um polipeptídeo semelhante à bHLH11 como definido neste.

O ácido nucleico pode ser introduzido diretamente em uma célula de planta ou na própria planta (incluindo introdução em um tecido, órgão ou qualquer outra parte de uma planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é preferivelmente introduzido em uma planta pela transformação. O termo "transformação" é descrito em mais detalhes na seção de "definições" neste.

As células de planta geneticamente modificadas podem ser regeneradas por intermédio de todos os métodos com que o trabalhador habilitado seja familiar. Os métodos adequados podem ser observados nas publicações anteriomente mencionadas por S.D. Kung and R. Wu, Potrykus or Hofgen and Willmitzer.

No geral após a transformação, as células de planta ou agrupamentos celulares são selecionados pela presença de um ou mais marcadores que são codificados pelos genes expressíveis das plantas co-transferidos com o gene de interesse, seguindo que o material transformado é regenerado em um total da planta. Para selecionar as plantas transformadas, o material vegetal obtido na transformação é, como uma regra, submetido às condições seletivas de modo que as plantas transformadas podem ser distinguidas das plantas transformadas. Por exemplo, as sementes obtidas na maneira descrita acima podem ser plantada e, após um período de crescimento inicial, submetido a uma seleção adequada pela pulverização. Uma possibilidade adicional consiste no crescimento das sementes, se apropriado após a esterilização, em placas ágar usando um agente de seleção adequado de modo que apenas as sementes transformadas podem desenvolver-se nas plantas. Alternativamente, as plantas transformadas são avaliadas pela

presença de um marcador selecionável tal como um descrito acima.

Seguindo a transferência e regeneração de DNA, as plantas putativamente transformadas também podem ser avaliadas, por exemplo usando a análise Southern, pela presença do gene de interesse, número de cópia e/ou organização genômica. Alternativamente ou adicionalmente, os níveis de expressão do DNA introduzido novamente pode ser monitorado usando análise Northern e/ou Western, ambas as técnicas sendo bem conhecidas às pessoas tendo habilidades comuns na técnica.

As plantas transformadas geradas pode ser propagadas por uma variedade de meios, tal como pela propagação clonal ou técnicas de geração clássica. Por exemplo, uma primeira geração da planta transformada (ou T1) pode ser por si mesma e transformantes de segunda geração de homozigoto (ou T2) selecionados, e as plantas T2 ainda podem ser então propagadas através de técnicas de geração clássicas. Os organismos transformados gerados podem tornar-se uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e células não transformadas; transformantes clonais (por exemplo, todas as células transformadas por conter o cassete de expressão); enxertos de tecidos transformados e não transformados (por exemplo, em plantas, um rizoma transformado enxertado em uma muda não transformada).

A presente invenção claramente estende-se a qualquer célula da planta ou planta produzida por qualquer um dos métodos descritos neste, e por todas partes da planta e propagules destes. A presente invenção estende-se ainda por abranger a progênie de uma célula transfectada ou transformada primária, tecido, órgão ou planta total que foi produzida por qualquer do métodos anteriomente mencionados, apenas o requerimento sendo que a progênie exibe a mesma características fenotípicas e/ou genotípicas como aquelas produzidas pelo padrão nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um

ácido nucleico isolado que codifica um polipeptídeo semelhante à bHLH11 como definido neste acima. As células hospedeiras preferidas de acordo com a invenção são células de planta. As plantas hospedeiras para os ácidos nucleicos ou o vetor usado nos métodos de acordo com a invenção, o cassete de expressão ou construção ou vetor são, à princípio, vantajosamente todas as plantas, que são capazes de sintetizar os polipeptídeos usados no método inventivo.

Os métodos da invenção são vantajosamente aplicáveis por qualquer planta. As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à super família Viridiplantae, em particular, plantas monocotiledônias e dicotiledônias incluindo forragem ou legumes de forragem, plantas ornamentais, lavouras de alimento, árvores e arbustos. De acordo com uma forma de realização preferida da presente invenção, a planta é uma planta de cultivo. Exemplos de plantas de cultivo incluem soja, girassol, canola, alfafa, colza, linhaça, algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta monocotiledônia. Exemplos de plantas monocotiledônias incluem cana de açúcar. Mais preferivelmente a planta é um cereal. Exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo, espécie de trigo, espelta, *secale*, *einkorn*, *teff*, *milo* e aveias.

A invenção também estende-se partes colhíveis de uma planta tal como, mas não limitado a sementes, folhas, frutos, flores, hastes, raízes, rizomas, tubérculos e bulbos, que as partes colhíveis compreendem um ácido nucleico recombinante que codifica um polipeptídeo semelhante à bHLH11. A invenção além disso relata os produtos derivados, preferivelmente diretamente derivados, a partir da parte coletável de uma tal planta, tal como grânulos e pós secos, óleo, gordura e ácidos graxos, amido ou proteínas.

De acordo com uma característica preferida da invenção, a expressão modulada é a expressão aumentada. Os métodos para aumentar a

expressão dos ácidos nucleicos ou genes ou produtos dos genes, são bem documentados na técnica e exemplos são fornecidos na seção de definições.

Como mencionado acima, um método preferido para modular a expressão de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 é pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante à bHLH11; entretanto os efeitos da realização do método, isto é itensificando os traços relacionados com o rendimento também podem ser atingido usando outras técnicas bem conhecidas, incluindo mas não limitado a rotulação de ativação de T-DNA, TILLING, recombinação homóloga. Uma descrição destas técnicas é fornecida na seção de definições.

A presente invenção também abrange o uso dos ácidos nucleicos que codificam os polipeptídeos semelhantes à bHLH11 as descrito neste e uso destes polipeptídeo semelhante à bHLH11s na itensificação de qualquer um dos traços mencionados acima relacionados com o rendimento em plantas.

Os ácidos nucleicos que codificam os polipeptídeos semelhante à bHLH11 descritos neste ou o polipeptídeo semelhante à bHLH11 por si mesmo, pode observar nos programas de geração que um marcador de DNA é identificado que pode ser geneticamente ligado a um gene que codifica o polipeptídeo semelhante à bHLH11. Os ácidos nucleicos/genes ou os polipeptídeos semelhantes à bHLH11 por si mesmos podem ser usados para definir um marcador molecular. Este marcador de DNA ou proteína então podem ser usados nos programas da geração para selecionar as plantas tendo traços intensificados relacionados com o rendimento como definido neste acima nos métodos da invenção.

As variantes alélicas de um ácido nucleico que codifica o polipeptídeo semelhante à bHLH11/gene também podem ser observado em programas de geração assistidos pelo marcador. Tais programas de geração

algumas vezes requerem a introdução da variação alélica pelo tratamento mutagênico das plantas, usando por exemplo a mutagênese EMS; alternativamente, o programa pode comerçar com uma coleção de variantes alélicas denominadas de origem "natural" causadas não itensionalmente. A identificação das variantes alélicas então acontece, por exemplo, por PCR. Este é seguido por uma etapa para a seleção de variantes alélicas superiores da sequência em questão e que dá rendimento aumentado. A seleção é tipicamente realizada pelo desempenho do desenvolimento do monitoramento de plantas contendo variantes alélicas diferentes da sequência em questão. O desempenho do desenvolvimento pode ser monitorado em uma estufa ou no campo. As etapas opcionais adicionais incluem o cruzamento de plantas que uma variante alélica superior foi indetificada com uma outra planta. Este deve ser usado, por exemplo, para fabricar uma combinação de características fenotípicas interessantes.

Os ácidos nucleicos que codificam os polipeptídeos semelhantes à bHLH11 também podem ser usados como sondas para o mapeamento físico e genético dos genes que estes estão a parte de, e como marcadores para traços ligados aqueles genes. Tal informação pode ser útil na geração de planta a fim de desenvolver as linhas com fenotipos desejados. Tal uso dos ácidos nucleicos que codificam o polipeptídeo semelhante à bHLH11 requer apenas uma sequência de ácido nucleico de pelo menos 15 nucleotídeos de comprimento. Os ácidos nucleicos que codificam o polipeptídeo semelhante à bHLH11 pode ser usados como polimorfismo de comprimento do fragmento de marcadores de restrição (RFLP). Southern blots (Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico da planta digerido por restrição pode ser sondado com os ácidos nucleicos semelhantes à bHLH11. Os padrões de grupos resultantes podem então ser submetidos à análise genética usando programas de computador tal como MapMaker (Lander et al. (1987)

Genomics 1: 174-181) a fim de construir um mapa genético. Além disso, os ácidos nucleicos podem ser usados para sondar Southern blots contendo DNA genômico tratado por endonuclease por restrição de uma série de indivíduos que representam o padrão e progênie de um cruzamento genético definido. A segregação do polimorfismo de DNA é notado e usado para calcular a posição do ácido nucleico que codifica o polipeptídeo semelhante à bHLH11 no mapa genético previamente obtido usando esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

A produção e uso de sondas derivadas do gene vegetal para o uso no mapeamento genético é descrito em Bernatzky and Tanksley (1986) Plant Mol. Biol. Reporter 4: 37-41. Numerosas publicações descrevem o mapeamento genético dos clones de cDNA específicos usando a metodologia resumida acima ou variações destas. Por exemplo, populações intercruzadas F2, populações cruzadas novamente, populações aleatoriamente auxiliares, linhas isogênicas próximas, e outras séries de indivíduos podem ser usados para o mapeamento. Tais metodologias são bem conhecidas àqueles habilitados na técnica.

As sondas do ácido nucleico também podem ser usadas para o mapeamento físico (isto é, disposição das sequências nos mapas físicos; ver Hoheisel et al. In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346, e referências citadas neste).

Em uma outra forma de realização, as sondas dos ácido nucleico podem ser usados na fluorescência direta no mapeamento da hibridização in situ (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora os métodos correntes do mapeamento FISH favorece o uso de amplos clones (diversos kb à diversas centenas de kb; ver Laan et al. (1995) Genome Res. 5:13-20), melhoramento na sensibilidade pode permitir o desempenho do mapeamento FISH usando sondas mais curtas.

Uma variedade da amplificação de ácido nucleico com base

em métodos para o mapeamento físico e genético pode ser realizado usando os ácidos nucleicos. Exemplos incluem a amplificação específica por alelos (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismo dos fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332), ligação específica por alelos (Landegren et al. (1988) Science 241:1077-1080), reações de extensão por nucleotídeos (Sokolov (1990) Nucleic Acid Res. 18:3671), mapeamento híbrido de radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e mapeamento Happy (Dear and Cook (1989) Nucleic Acid Res. 17:6795-6807). Para estes métodos, a sequência de um ácido nucleico é usada para projetar e produzir pares de iniciadores para o uso na reação de amplificação ou nas reações da extensão iniciadora. O projeto de tais iniciadores é bem conhecido àqueles habilitados na técnica. Em métodos que utilizam o mapeamento genético com base em PCR, este pode ser necessário para identificar as diferenças da sequência de DNA entre os padrões do cruzamento do mapeamento na região correspondente à sequência do ácido nucleico imediata. Este, entretanto, não é necessário no geral para os métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo traços intensificados relacionados com o rendimento, como descrito anteriormente acima. Estes traços também podem ser combinados com outros traços economicamente vantajosos, tal como traços que itensificam a produção adicional, tolerância a outras tensões bióticas e abióticas, traços que modificam várias características arquiteturais e/ou características bioquímicas e/ou fisiológicas.

VI. Polipeptídeo ASR (ácido abscísico-, tensão-, e maturativo induzido)

Um método preferido para modular (preferivelmente, aumentar) a expressão de um ácido nucleico que codifica um polipeptídeo ASR é pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo ASR.

Qualquer referência à seguir a uma "proteína útil nos métodos da invenção" é determinada entender um polipeptídeo ASR como definido neste. Qualquer referência à seguir a um "ácido nucleico útil nos métodos da invenção" é determinada entender um ácido nucleico capaz de codificar um tal polipeptídeo ASR. O ácido nucleico a ser introduzido em uma planta (e, portanto, útil na realização dos métodos da invenção) é qualquer ácido nucleico que codifica o tipo de proteína que será descrito agora, à seguir também denominado "ácido nucleico ASR" ou "gene ASR".

Um "polipeptídeo ASR" como definido neste refere-se às proteínas representadas pela SEQ ID N°: 397 e pelos homólogos (ortólogos e parálogos) destes.

Preferivelmente, os homólogos da SEQ ID N°: 397 tem um domínio ABA WDS (entrada Pfam PF02496).

Preferivelmente os polipeptídeos ASR adicionais da invenção são aqueles tendo a ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 83 %, 85 %, 87 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, ou mais identidade de sequência a qualquer um dos polipeptídeos dados na Tabela F1.

O termo "domínio" e "motivo" é definido na seção de "definições" neste. Os bancos de dados especializados existem para a identificação de domínios, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-244), InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318), Prosite (Bucher and Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (In) ISMB-94; Proceedings 2° International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004)) ou Pfam (Bateman et al., Nucleic

Acids Research 30(1): 276-280 (2002)). Uma série de ferramentas para a análise *in silico* de sequências de proteínas está disponível no servidor proteômico ExPASy (Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31:3784-3788(2003)). Os domínios ou motivos também podem ser identificados usando técnicas de rotina, tal como pelo alinhamento da sequência.

Os métodos para o alinhamento de sequências para comparação são bem conhecidos na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. GAP usa o algoritmo de Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar o alinhamento global (isto é medindo-se as sequências completas) de duas sequências que maximizam o número de equiparações e minimiza o número de fendas. O algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-10) calcula a identidade de sequência percentual e realiza uma análise estatística da similaridade entre as duas sequências. O software para realizar análise BLAST está publicamente disponível através do National Centre for Biotechnology Information (NCBI). Os homólogos podem ser facilmente identificados usando-se, por exemplo, o algoritmo de alinhamento de sequência múltipla ClustalW (versão 1.83), com os parâmetros de alinhamento default e um método de registro em porcentagem. As porcentagens globais de similaridade e de identidade também podem ser determinadas usando-se um dos métodos disponíveis no pacote de software MatGAT (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando-se proteína ou sequências de DNA.). A edição manual menor pode ser realizada para otimizar o alinhamento entre os motivos conservados, como seria aparente à pessoa habilitada na técnica. Além disso, em vez de usar as sequências de comprimento total para a identificação de homólogos, domínios

específicos também podem ser usados. Os valores de identidade de sequência podem ser determinados sobre o aminoácido inteiro ou sequência de aminoácido ou em domínios selecionados ou motivos conservados, usando-se os programas mencionados acima usando-se os parâmetros default.

Além disso, os polipeptídeos ASR (pelo menos em sua forma natural), tão perto quanto da SEQ ID N°: 397 e estes homólogos são interessados, tipicamente têm a capacidade para aumentar a resistência da tensão salina de plantas. As ferramentas e técnicas para a expressão de ASR em plantas e teste para a resistência da tensão salina aumentada são bem conhecidos na técnica.

A presente invenção é ilustrada transformando-se as plantas com a sequência de ácido nucleico representado pela SEQ ID N°: 396, que codifica a sequência de polipeptídeo da SEQ ID N°: 397. Entretanto, desempenho da invenção não está restrito a estas sequências; os métodos da invenção podem ser vantajosamente realizados usando-se qualquer ácido nucleico que codifica ASR ou polipeptídeo ASR como definido neste.

Exemplos de ácidos nucleicos que codificam os polipeptídeos ASR podem ser observados em banco de dados conhecidos na técnica. Tais ácidos nucleicos são úteis na realização dos métodos da invenção. Ortólogos e parálogos, os termos "ortólogos" e "parálogos" sendo como definido neste, já podem ser identificados pela realização de uma denominada pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envolve o submetro de uma sequência em dúvida ao BLAST (por exemplo usando SEQ ID N°: 397) contra qualquer banco de dados de sequência, tal como os bancos de dados NCBI publicamente disponíveis. BLASTN ou TBLASTX (usando-se os valores default padrão) são, no geral, usados quando começa-se a partir de uma sequência de nucleotídeo, e BLASTP ou TBLASTN (usando-se os valores default padrão) quando começa-se a partir de uma sequência de proteína. Os resultados de BLAST podem ser opcionalmente filtrados. As

sequências de comprimento total de cada um dos resultados filtrados ou resultados não filtrados são então submetidos ao BLAST novamente (segundo BLAST) contra as sequências do organismo a partir do qual a sequência em questão é derivada (quando a sequência em questão é SEQ ID N°: 396 ou SEQ ID N°: 397, o segundo BLAST portanto, deve ser contra as sequências de *Oryza sativa*). Os resultados do primeiro e segundo BLAST são então comparados. Um parálogo é identificado se um acerto de classificação alta do primeiro blast da mesma espécie como a partir da qual a sequência em questão é derivada, submetendo-se ao BLAST novamente então resulta idealmente na sequência em questão entre os acertos mais altos; um ortólogo é identificado se um acerto de classificação alta no primeiro BLAST não da mesma espécie como a partir da qual a sequência em questão é derivada e preferivelmente resulta no BLAST novamente na sequência em questão estando entre os acertos mais altos.

Os acertos de classificação alta são aqueles tendo um valor E baixo. Quanto mais baixo o valor E, mais significante o registro (ou, em outras palavras, menor a chance que o acerto seja encontrado por chance).

A computação do valor E é bem conhecido na técnica. Além dos valores E, as comparações são registradas pela identidade de porcentagem. A identidade de porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácidos nucleicos comparadas (ou polipeptídeo) em um comprimento particular. No caso de famílias grandes, ClustalW pode ser usado, seguido por uma árvore de união próxima, para ajudar a visualizar o agrupamento de genes relacionados e para identificar ortólogos e parálogos.

As variantes de ácido nucleico que codifica os homólogos e derivados da SEQ ID N°: 397 também podem ser úteis na prática dos métodos da invenção, os termos "homólogo" e "derivado" sendo como definido neste. Também são úteis nos métodos da invenção os ácidos nucleicos que

codificam homólogos e derivados de ortólogos ou parálogos da SEQ ID N°: 397. Homólogos e derivados úteis nos métodos da presente invenção têm substancialmente a mesma atividade biológica e funcional como a proteína não modificada da qual estes são derivados.

As variantes de ácido nucleico adicionais úteis na prática dos métodos da invenção incluem as porções de ácidos nucleicos que codificam os polipeptídeos ASR, ácidos nucleicos que hibridizam-se a ácidos nucleicos que codificam os polipeptídeos ASR, variantes de união de ácidos nucleicos que codificam os polipeptídeos ASR, variantes alélicas de ácidos nucleicos que codificam os polipeptídeos ASR e variantes de ácidos nucleicos que codificam os polipeptídeos ASR obtidos pelo embaralhamento de gene. Os termos sequência de hibridização, variante de união, variante alélica e embaralhamento de gene são como descritos neste.

Os ácidos nucleicos que codificam os polipeptídeos ASR não necessitam ser ácidos nucleicos de comprimento total, visto que o desempenho dos métodos da invenção não contam com o uso de sequências de comprimento total de ácido nucleico. De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta uma porção da SEQ ID N°: 396 ou uma porção de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo da SEQ ID N°: 397.

Uma porção de um ácido nucleico pode ser preparada, por exemplo, realizando-se uma ou mais anulações no ácido nucleico. As porções podem ser usadas na forma isolada ou estas podem ser fundidas a outras sequências codificadoras (ou não codificadoras) a fim de, por exemplo, produzir uma proteína que combine diversas atividades. Quando fundido a outras sequências codificadoras, o polipeptídeo resultante produzido na tradução pode ser maior do que aquele predito para a porção de proteína.

As porções úteis nos métodos da invenção, codificam um

polipeptídeo ASR como definido nestee têm substancialmente a mesma atividade biológica como as sequências de aminoácido dada na SEQ ID N°: 397. Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos dada na SEQ ID N°: 396 ou é uma porção de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das sequências de aminoácido dada na SEQ ID N°: 396. Preferivelmente a porção é pelo menos 400, 450, 500, 550, 600, 650, 700, 750 nucleotídeos consecutivos de comprimento, os nucleotídeos consecutivos sendo da SEQ ID N°: 396 ou de um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 397. Mais preferivelmente a porção é uma porção do ácido nucleico da SEQ ID N°: 396.

Uma outra variante de ácido nucleico útil nos métodos da invenção é um ácido nucleico capaz de hibridizar, sob condições de estringência reduzida, preferivelmente sob condições estringentes, com um ácido nucleico que codifica um polipeptídeo ASR como definido neste ou com uma porção como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta, um ácido nucleico capaz de hibridizar à SEQ ID N°: 396 ou que compreende introduzir e expressar em uma planta, um ácido nucleico capaz de hibridizar a um ácido nucleico que codifica um ortólogo, parálogo ou homólogo da SEQ ID N°: 396.

As sequências de hibridização úteis nos métodos da invenção codificam um polipeptídeo ASR como definido neste, tendo substancialmente a mesma atividade biológica como as sequências de aminoácido dada na SEQ ID N°: 397. Preferivelmente, a sequência de hibridização é capaz de hidridizar à SEQ ID N°: 396 ou a uma porção de qualquer uma destas sequências, uma porção sendo como definido acima ou a sequência de hibridização é capaz de hibridizar a um ácido nucleico que codifica um

ortólogo ou parálogo da SEQ ID N°: 397.

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante de união que codifica um polipeptídeo ASR como definido neste acima, a variante de união sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta de uma variante de união da SEQ ID N°: 396 ou a variante de união de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo da SEQ ID N°: 397.

Uma outra variante de ácido nucleico útil na realização dos métodos da invenção é uma variante alélica de um ácido nucleico que codifica um polipeptídeo ASR como definido neste acima, uma variante alélica sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, que compreende introduzir e expressar em uma planta de uma variante alélica da SEQ ID N°: 396 ou que compreende introduzir e expressar em uma planta de uma variante alélica de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo das sequências de aminoácido representado pela SEQ ID N°: 397.

As variantes alélicas úteis nos métodos da presente invenção tem substancialmente a mesma atividade biológica como o polipeptídeo ASR da SEQ ID N°: 397. As variantes alélicas existem na natureza, e abrange dentro dos métodos da presente invenção é o uso destes alelos naturais. A mistura do gene ou evolução direcionada também pode ser usada para gerar variantes dos ácidos nucleicos que codificam os polipeptídeos ASR como definido acima; o termo "a mistura do gene" sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar os traços relacionados com o rendimento em plantas, que

compreende introduzir e expressar em uma planta de uma variante da SEQ ID N°: 396 ou que compreende introduzir e expressar em uma planta de uma variante de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo da SEQ ID N°: 397, que a variante do ácido nucleico é obtido pelo embaralhamento de gene.

Além disso, as variantes do ácido nucleico também podem ser obtidas pela mutagênese direcionada ao local. Diversos métodos são disponíveis para atingir a mutagênese direcionada ao local, o mais comum sendo PCR com base em métodos (Current Protocols in Molecular Biology. Wiley Eds.).

Os ácidos nucleicos que codificam os polipeptídeos ASR podem ser derivados a partir de qualquer fonte artificial ou natural. O ácido nucleico pode ser modificada a partir da forma natural na composição e/ou ambiente genômico através da consideração manipulação humana. Preferivelmente o ácido nucleico que codifica o polipeptídeo ASR é a partir de uma planta. No caso da SEQ ID N°: 396, o ácido nucleico que codifica o polipeptídeo ASR é preferivelmente de uma planta monocotiledônea, mais preferivelmente da família Poaceae, mais preferivelmente o ácido nucleico é de *Oryza sativa.*

A invenção também fornece até agora ácidos nucleicos que codificam ASR não conhecidos e polipeptídeos ASR.

De acordo com uma outra forma de realização da presente invenção, portanto, é fornecido uma molécula de ácido nucleico isolado selecionado de:

(i) um ácido nucleico representado por qualquer uma da SEQ ID N°: 401, 403, 405, 407, 409, 411, 413, 415 e 417;

(ii) o complemento de um ácido nucleico representado por qualquer uma da SEQ ID N°: 401, 403, 405, 407, 409, 411, 413, 415 e 417;

(iii) um ácido nucleico que codifica um polipeptídeo ASR

tendo, na ordem crescente de preferência, pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência à sequência de aminoácido representada por qualquer uma da SEQ ID N°: 402, 404, 406, 408, 410, 412, 414, 416 e 418.

De acordo com uma outra forma de realização da presente invenção, também existe fornecido um polipeptídeo isolado selecionado de:

(i) uma sequência de aminoácido representada por qualquer uma da SEQ ID N°: 402, 404, 406, 408, 410, 412, 414, 416 e 418;

(ii) uma sequência de aminoácido tendo, na ordem crescente de preferência, pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência à sequência de aminoácido representado por qualquer uma da SEQ ID N°: 402, 404, 406, 408, 410, 412, 414, 416 e 418.

(iii) derivados de qualquer uma das sequências de aminoácido dadas em (i) ou (ii) acima.

O desempenho dos métodos da invenção dá plantas tendo traços intensificados relacionados com o rendimento. Em particular, o desempenho dos métodos da invenção dá plantas tendo vigor precoce aumentado e rendimento aumentado, especialmente biomassa aumentada e rendimento aumentado de semente com relação às plantas de controle. Os termos "produção" e "rendimento de semente" são descritos em mais detalhes na seção de "definições" neste.

A referência neste aos traços intensificados relacionados com o rendimento é determinada entender um aumento no vigor precoce e/ou na biomassa (peso) por uma ou mais partes de uma planta, que pode incluir partes acima do solo (coletáveis) e/ou partes abaixo do solo (coletáveis). Em particular, tais partes colhíveis são biomassa e/ou sementes, e o desempenho dos métodos da invenção resultam em plantas tendo o vigor precoce aumentado, biomassa e/ou rendimento de semente relativo ao vigor precoce,

biomassa ou rendimento de semente de plantas de controle.

Tomando o milho como um exemplo, um aumento da produção pode ser manifestada como um ou mais do seguinte: aumento no número de plantas estabelecidas por hectare ou acre, um aumento no número de espigas por planta, um aumento no número de séries, número de grãos por série, peso dos grãos, peso dos grãos milhar, comprimento/diâmetro da espiga, aumento da taxa de enchimento da semente (que é o número de sementes enchidas dividida pelo número total de sementes e multiplicadas por 100), entre outros. Tomando o arroz como um exemplo, um aumento da produção pode manifestar por si mesmo como um aumento em um ou mais do seguinte: número de plantas por hectare ou acre, número de panículas por planta, número de espigas por panículas, número de flores (florzinhas) por panícula (que é expressado como uma razão do número de sementes enchidas no número de panícula primárias), aumento da taxa de enchimento da semente (que é o número de sementes enchidas dividida pelo número total de sementes e multiplicadas por 100), aumento no peso dos grãos milhar, entre outros.

A presente invenção fornece um método para aumentar a produção, especialmente biomassa e/ou rendimento de semente de plantas, com relação às plantas de controle, que o método compreende a expressão modular, preferivelmente aumentar a expressão, em uma planta de um ácido nucleico que codifica um polipeptídeo ASR como definido neste.

Visto que as plantas transgênicas de acordo com a presente invenção tem o rendimento aumentado, é igualmente que estas plantas exibam uma taxa de desenvolvimento aumentado (durante pelo menos parte de seu ciclo de vida), relativo à taxa de desenvolvimento de plantas de controle em um estágio correspondente no seu ciclo de vida.

A taxa de desenvolvimento aumentada pode ser específica por uma ou mais partes de uma planta (incluindo sementes) ou pode estar em toda parte substancialmente de uma planta total. As plantas tendo uma taxa de

desenvolvimento aumentado pode ter um ciclo de vida mais curto. O ciclo de vida de uma planta pode ser levado significar um período necessário para o desenvolvimento a partir da semente madura seca até o estágio onde a planta tem produzido as sementes maduras secas, similar ao material de partida. Este ciclo de vida pode ser influenciado pelos fatores tal como vigor precoce, taxa de desenvolvimento, índice de viço, período e velocidade de florescência do amadurecimento da semente. O aumento na taxa de desenvolvimento pode acontecer em um ou mais estágios no ciclo de vida de uma planta ou durante substancialmente ao ciclo de vida total da planta. A taxa aumentada de desenvolvimento durante os estágios precoces no ciclo de vida de uma planta pode refletir o vigor itensificado. O aumento na taxa de desenvolvimento pode alterar o ciclo de coleta de uma planta permitindo as plantas serem semeadas depois e/ou mais cedo coletadas do que de outra maneira seria possível (um efeito similar pode ser obtido com período de florescência precoce). Se a taxa de desenvolvimento é suficientemente aumentada, este pode deixar para a semeadura adicional de sementes das mesmas espécies de plantas (por exemplo semeadura e coletagem de plantas de arroz seguido pela semeadura e coletagem das plantas de arroz adicionais todas dentro de um período de desenvolvimento convencional). Similarmente, se a taxa de desenvolvimento é suficientemente aumentada, este pode deixar para a semeadura adicional de sementes de espécies de plantas diferentes (por exemplo a semeadura e coletagem de plantas de milho seguido pela, por exemplo, semeadura e coletagem opcional de soja, batata ou qualquer outra planta adequada). Os períodos de coletagem adicional do mesmo rizoma no caso de algumas plantas de cultivo também podem ser possível. Alterando o ciclo de coleta de uma planta pode levar a um aumento na produção da biomassa anual por acre (devido a um aumento no número de vezes (dito em um ano) que qualquer planta particular pode ser desenvolvida e coletada). Um aumento na taxa de desenvolvimento também pode permitir o cultivo de

plantas transgênicas em uma ampla área geográfica do que suas contrapartes de tipo selvagem, visto as limitações territoriais para o desenvolvimento de uma lavoura são frequentemente determinados pelas condições ambientais adversas no período do plantio (temporada precoce) ou no período de coletagem (temporada tardia). Tais condições adversas podem ser evitadas se o ciclo de coleta seja mais curto. A taxa de desenvolvimento pode ser determinação pela derivação de vários parâmetros a partir de curvas de desenvolvimento, tais parâmetros podem ser: T-Mid (o período de retirada para as plantas atingirem 50 % do seu tamanho máximo) e T-90 (período de retirada para as plantas atingirem 90 % do seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção dá plantas tendo uma taxa de desenvolvimento aumentado com relação às plantas de controle. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a taxa de desenvolvimento de plantas, que o método compreende a expressão modular, preferivelmente aumentar expression, em uma planta de um ácido nucleico que codifica um polipeptídeo ASR como definido neste. Em uma forma de realização particular, o desempenho dos métodos da presente invenção dá plantas com o vigor precoce aumentado.

Um aumento na produção e/ou taxa de desenvolvimento ocorre se a planta está sob condições sem estresse ou se a planta está exposta a várias tensões em comparação à plantas de controle. As plantas tipicamente respondem à exposição à tensão pelo crescimento mais lentamente. Em condições de diversas tensões, a planta ainda pode interromper o desenvolvimento completamente. A tensão branda de outra maneira é definida neste como sendo qualquer tensão que qual a planta é exposta que não resulta no interrompimento da planta para desenvolver completamente sem a capacidade de recuperar o desenvolvimento. A tensão branda no sentido da

invenção leva a uma redução no desenvolvimento das plantas submetidas à tensão de menos do que 40 %, 35 % ou 30 %, preferivelmente menos do que 25 %, 20 % ou 15 %, mais preferivelmente menos do que 14 %, 13 %, 12 %, 11 % ou 10 % ou menos em comparação à planta de controle sob condições sem estresse. Devido as vantagens nas práticas da agricultura (irrigação, fertilização, tratamentos com pesticidas) diversas tensões não são frequentemente encontradas em plantas de cultivos cultivadas. Como uma consequência, o desenvolvimento compromissado induzido pela tensão branda é frequentemente uma característica indesejável para a agricultura. As tensões brandas são as tensões abióticas e/ou bióticas (ambientais) que a planta é exposta. As tensões abióticas podem ser devido à aridez ou excesso de água, tensão anaeróbica, tensão salina, toxicidade química, tensão oxidativa e temperaturas de frio, calor e congelamento. A tensão abiótica pode ser uma tensão osmótica causa pela tensão à água (particularmente devido à aridez), tensão salina, tensão oxidativa ou uma tensão iônica. As tensões bióticas são tipicamente aquelas tensões causadas pelos patogênios, tal como bactérias, vírus, fungos e insetos.

Em particular, os métodos da presente invenção podem ser realizados sob condições sem estresse ou sob condições de aridez branda para dar plantas tendo rendimento aumentado com relação às plantas de controle. Como relatado em Wang et al. (Planta (2003) 218: 1-14), tensão abiótica leva a uma sére de mudanças moleculares, bioquímicas, fisiológicas e morfológicas que afetam adversamente o desenvolvimento da planta e produtividade. A aridez, salinidade, temperaturas extremas e tensão oxidativa estão conhecidas serem interconectadas e podem induzir o desenvolvimento e dano celular através de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve particularmente um alto grau de "cruzamento" entre a tensão de aridez e alta tensão por salinidade. Por exemplo, aridez e/ou salinização são manifestados principalmente como

tensão osmótica, resultando no rompimento da homeostase e distribuição de íon na célula. A tensão oxidativa, que frequentemente acompanha temperatura alta ou baixa, salinidade ou tensão de aridez, pode causar a desnaturação de proteínas estruturais e funcionais. Como uma consequência, estas diversas tensões ambientais frequentemente ativam os caminhos de sinalização celular similares e respostas celulares, tal como a produção de proteínas de tensão, super regulação de anti-oxidantes, acúmulo de solutos compatíveis e suspensão do desenvolvimento. O termo de condições de "não tensão" como usado neste são aquelas condições ambientais que permitem o ótimo desenvolvimento das plantas. As pessoas habilitadas na técnica são atentas das condições do solo normal e condições climáticas para uma localização dada. As plantas com ótimas condições de desenvolvimento, (desenvolvimento sob condições sem estresse) tipicamente produzem na ordem crescente de preferência pelo menos 90 %, 87 %, 85 %, 83 %, 80 %, 77 % ou 75 % da produção média de tal planta em um dado ambiente. A produção média pode ser calculada na base da coleta e/ou temporada. As pessoas habilitadas na técnica estão atentas das produções de rendimento médio de uma lavoura.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições sem estresse ou sob condições de secura brandas rendimento aumentado e/ou vigor precoce aumentado, com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a produção e/ou vigor precoce em plantas desenvolvidas sob condições sem estresse ou sob condições de secura brandas, que o método compreende aumentar expressão em uma planta de um ácido nucleico que codifica um polipeptídeo ASR.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições de deficiência do nutriente, particularmente sob

condições de deficiência de nitrogênio, rendimento aumentado com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a produção em plantas desenvolvidas sob condições de deficiência do nutriente, que o método compreende aumentar expressão em uma planta de um ácido nucleico que codifica um polipeptídeo ASR. Deficiência do nutriente pode resultar a partir da perda de nutrientes tal como nitrogênio, fosfatos e outros compostos contendo fósforo, potássio, cálcio, cadmío, magnésio, manganês, ferro e boro, entre outros.

A presente invenção abrange as plantas ou partes destas (incluindo sementes) obtidas pelos métodos de acordo com a presente invenção. As plantas ou partes destas compreende um transgene de ácido nucleico que codifica um polipeptídeo ASR como definido acima.

A invenção também fornece as construções genéticas e vetores para facilitar a introdução e/ou expressão em plantas dos ácidos nucleicos que codificam os polipeptídeos ASR. A construção de genes pode ser inserida nos vetores, que pode ser comercialmente disponível, adequado para a transformação em plantas e adequado para a expressão do gene de interesse nas células transformadas. A invenção também fornece o uso de uma construção de gene como definido neste nos métodos da invenção.

Mais especificamente, a presente invenção fornece uma construção que compreende:

(a) um ácido nucleico que codifica um polipeptídeo ASR como definido acima;

(b) uma ou mais sequências de controle capazes de conduzir a expressão da sequência de ácido nucleico de (a) e opcionalmente

(c) uma sequência de terminação de transcrição.

Preferivelmente, o ácido nucleico que codifica um polipeptídeo ASR é como definido acima. O termo "sequência de controle" e

"sequência de terminação" são como definidos neste.

As plantas são transformadas com um vetor que compreende qualquer um dos ácidos nucleicos descritos acima. O técnico habilitado é bem atento aos elementos genéticos que devem estar presentes no vetor a fim de transformar sucessivamente, selecionar e propagar as células hospedeiras contendo a sequência de interesse. A sequência de interesse é operacionalmente ligada a uma ou mais sequências de controle (pelo menos a um promotor).

Vantajosamente, qualquer tipo de promotor, se natural ou sintético, pode ser usado para conduzir a expressão de uma sequência de ácido nucleico. Um promotor constitutivo é particularmente útil nos métodos da invenção. Ver a seção de "Definições" neste pelas definições de vários tipos de promotores.

Deve ser claro que a aplicabilidade da presente invenção não é restrito ao ácido nucleico que codifica o polipeptídeo ASR representado pela SEQ ID N°: 396, ou é a aplicabilidade da invenção restrita à expressão de um ácido nucleico que codifica o polipeptídeo ASR quando conduzido por um promotor constitutivo específico.

O promotor constitutivo é preferivelmente um promotor GOS2, preferivelmente um promotor GOS2 de arroz. Ainda preferivelmente o promotor constitutivo é representado por uma sequência de ácido nucleico substancialmente similar à SEQ ID N°: 398, mais preferivelmente o promotor constitutivo é como representado pela SEQ ID N°: 398. Ver a seção das "Definições" neste pelos exemplos adicionais de promotores constitutivos.

Opcionalmente, uma ou mais sequências terminadoras podem ser usadas na construção introduzida em uma planta. Os elementos reguladores adicionais podem incluir a transcrição bem como os itensificadores de tradução. Aqueles habilitados na técnica serão atentos das sequências itensificadoras e terminadoras que pode ser adequado para o uso

na realização da invenção. Uma sequência de íntron também podem ser adicionada à região de não tradução da extremidade final 5' (UTR) ou na sequência codificadora para aumentar a quantidade da mensagem madura que acumula no citosol, como descrito na seção de definições. Outras sequências de controles (perto do promotor, itensificador, silenciador, sequências de íntron, região 3' UTR e/ou 5' UTR) podem ser elementos de estabilização de RNA e/ou proteína. Tais sequências seria conhecida ou já pode ser obtida por uma pessoa habilitada na técnica.

As construções genéticas da invenção ainda podem incluir uma origem da sequência de replicação que é requerida pela manutenção e/ou reprodução em um tipo celular específico. Um exemplo é quando uma construção genética é requerida ser mantida em uma células bacteriana como um elemento genético episssomal (por exemplo, moléculas de plasmídio e cosmídio). As origens preferidas da reprodução incluem, mas não são limitadas a, o f1-ori e colE1.

Para a detecção da transferência sucessível das sequências de ácido nucleico como usado nos métodos da invenção e/ou seleção de plantas transgênicas que compreendem estes ácidos nucleicos, é vantajoso para o uso dos genes marcadores (ou genes repórter). Portanto, a construção genética pode opcionalmente compreender um gene de marcador selecionável. Os marcadores selecionáveis são descritos em mais detalhes na seção de "definições" neste. Os genes marcadores podem ser removidos ou taxados a partir da célula transgênica uma vez que estes não são necessariamente mais longos. As técnicas para a remoção do marcador são conhecidas na técnica, as técnicas úteis são descritas acima na seção de definições.

A invenção também fornece um método para a produção de plantas transgênicas tendo traços intensificados relacionados com o rendimento com relação às plantas de controle, que compreende introdução e a expressão em uma planta de qualquer ácido nucleico que codifica um

polipeptídeo ASR como definido neste acima.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo traços intensificados aumentados relacionados com o rendimento, particularmente o vigor precoce aumentado e/ou rendimento aumentado, que o método compreende:

(i) introduzir e expressar em uma planta ou célula de planta de um ácido nucleico que codifica o polipeptídeo ASR; e

(ii) cultivar a célula de planta sob condições que promovam o crescimento e desenvolvimento de planta.

O ácido nucleico de (i) pode ser qualquer um dos ácidos nucleicos capaz de codificar um polipeptídeo ASR como definido neste.

O ácido nucleico pode ser introduzido diretamente em uma célula de planta ou na própria planta (incluindo introdução em um tecido, órgão ou qualquer outra parte de uma planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é preferivelmente introduzido em uma planta pela transformação. O termo "transformação" é descrito em mais detalhes na seção das "definições" neste.

As células de planta geneticamente modificadas podem ser regeneradas por intermédio de todos os métodos com que o trabalhador habilitado seja familiar. Os métodos adequados podem ser observados nas publicações anteriomente mencionadas por S.D. Kung and R. Wu, Potrykus or Hofgen and Willmitzer.

No geral após a transformação, as células de planta ou agrupamentos celulares são selecionadas pela presença de um ou mais marcadores que são codificados pelos genes expressíveis das plantas co-transferidos com o gene de interesse, seguindo que o material transformado é regenerado em um total da planta. Para selecionar as plantas transformadas, o material vegetal obtido na transformação é, como uma regra, submetido às condições seletivas de modo que as plantas transformadas podem ser

distinguidas das plantas transformadas. Por exemplo, as sementes obtidas na maneira descrita acima podem ser plantada e, após um período de crescimento inicial, submetido a uma seleção adequada pela pulverização. Uma possibilidade adicional consiste no crescimento das sementes, se apropriado após a esterilização, em placas ágar usando um agente de seleção adequado de modo que apenas as sementes transformadas podem desenvolver-se nas plantas. Alternativamente, as plantas transformadas são avaliadas pela presença de um marcador selecionável tal como um descrito acima.

Seguindo a transferência e regeneração de DNA, as plantas putativamente transformadas também podem ser avaliadas, por exemplo usando a análise Southern, pela presença do gene de interesse, número de cópia e/ou organização genômica. Alternativamente ou adicionalmente, os níveis de expressão do DNA introduzido novamente pode ser monitorado usando a análise Northern e/ou Western, ambas as técnicas sendo bem conhecidas às pessoas tendo habilidades comuns na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de meios, tal como pela propagação clonal ou técnicas de geração clássica. Por exemplo, uma primeira geração de planta transformada (ou T1) pode ser por si mesma e transformantes da segunda geração de homozigoto (ou T2) selecionados, e as plantas T2 ainda podem ser então propagadas através de técnicas de geração clássicas. Os organismos transformados gerados podem tornar-se uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e células não transformadas; transformantes clonais (por exemplo, todas as células transformadas por conter o cassete de expressão); enxertos de tecidos transformados e não transformados (por exemplo, em plantas, um rizoma transformado enxertado em uma muda não transformada).

A presente invenção claramente estende-se a qualquer célula da planta ou planta produzida por qualquer um dos métodos descrito neste, e

por todas partes da planta e propagules destes. A presente invenção estende-se ainda por abranger a progênie de uma célula transfectada ou transformada primária, tecido, órgão ou planta total que foi produzida por qualquer do métodos anteriomente mencionados, apenas o requerimento sendo que a progênie exibe a mesma características fenotípicas e/ou genotípicas como aquelas produzidas pelo padrão nos métodos de acordo com a invenção.

A invenção também inclui as células hospedeiras contendo um ácido nucleico isolado que codifica um polipeptídeo ASR como definido neste acima. As células hospedeiras preferidas de acordo com a invenção são células de planta. As plantas hospedeiras pelos ácidos nucleicos ou o vetor usados no método de acordo com a invenção, o cassete de expressão ou construção ou vetor são, à princípio, vantajosamente todas as plantas, que são capazes de sintetizar os polipeptídeos usados no método inventivo.

Os métodos da invenção são vantajosamente aplicáveis por qualquer planta. As plantas que são particularmente útil nos métodos da invenção incluem todas as plantas que pertencem à super família Viridiplantae, em particular, plantas monocotiledônias e dicotiledônias incluindo forragem ou legumes de forragem, plantas ornamentais, lavouras de alimento, árvores e arbustos. De acordo com uma forma de realização preferida da presente invenção, a planta é uma planta de cultivo. Exemplos de plantas de cultivo incluem soja, girassol, canola, alfafa, colza, algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta monocotiledônia. Exemplos de plantas monocotiledônias incluem cana de açúcar. Mais preferivelmente a planta é um cereal. Exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

A invenção também estende-se as partes colhíveis de uma planta tal como, mas não limitado a sementes, folhas, frutos, flores, hastes, raízes, rizomas, tubérculos e bulbos. A invenção além disso relata os produtos derivados, preferivelmente diretamente derivados, a partir da parte coletável

de uma tal planta, tal como grânulos e pós secos, óleo, gordura e ácidos graxos, amido ou proteínas.

De acordo com uma característica preferida da invenção, a expressão modulada é a expressão aumentada. Os métodos para aumentar a expressão de ácidos nucleicos ou genes ou produtos dos genes, são bem documentados na técnica e exemplos são fornecidos na seção de definições.

Como mencionado acima, um método preferido para modular (preferivelmente, aumentar) a expressão de um ácido nucleico que codifica um polipeptídeo ASR é pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo ASR; entretanto os efeitos da realização do método, isto é itensificando os traços relacionados com o rendimento também podem ser atingido usando outras técnicas bem conhecidas, incluindo mas não limitado a rotulação de ativação de T-DNA, TILLING, recombinação homóloga. Uma descrição destas técnicas é fornecida na seção de definições.

A presente invenção também abrange o uso dos ácidos nucleicos que codificam os polipeptídeos ASR como descrito neste e uso destes polipeptídeos ASR na itensificação de qualquer um dos traços mencionados acima relacionados com o rendimento em plantas.

Os ácidos nucleicos que codificam o polipeptídeo ASR descrito neste ou os polipeptídeos ASR por si mesmos, podem observar nos programas de geração em que um marcador de DNA é identificado que pode ser geneticamente ligado a um gene que codifica o polipeptídeo ASR. Os ácidos nucleicos/genes ou os polipeptídeos ASR por si mesmos podem ser usados para definir um marcador molecular. Este marcador de DNA ou proteína então pode ser usado nos programas da geração para selecionar as plantas tendo traços intensificados relacionados com o rendimento como definido neste acima nos métodos da invenção.

As variantes alélicas de um ácido nucleico/gene que codifica o

polipeptídeo ASR também podem ser observadas em programas de geração assistidas pelo marcador. Tais programas de geração algumas vezes requerem a introdução da variação alélica pelo tratamento de mutagênese das plantas, usando por exemplo a mutagênese EMS; alternativamente, o programa pode comerçar com uma coleção de variantes alélicas denominadas de origem "natural" causadas não itensionalmente. A identificação das variantes alélicas então acontece, por exemplo, por PCR. Este é seguido por uma etapa para a seleção de variantes alélicas superiores da sequência em questão e que dá rendimento aumentado. A seleção é tipicamente realizada pelo desempenho do desenvolimento do monitoramento de plantas contendo variantes alélicas diferentes da sequência em questão. O desempenho do desenvolvimento pode ser monitorado em uma estufa ou no campo. As etapas opcionais adicionais incluem o cruzamento de plantas em que a variante alélica superior foi indetificada com uma outra planta. Esta deve ser usada, por exemplo, para fabricar uma combinação de características fenotípicas interessantes.

Os ácidos nucleicos que codificam os polipeptídeos ASR também podem ser usados como sondas para o mapeamento físico e genético dos genes que estes estão a parte de, e como marcadores para traços ligados aqueles genes. Tal informação pode ser útil na geração de planta a fim de desenvolver as linhas com os fenotipos desejados. Tal uso dos ácidos nucleicos que codificam os polipeptídeos ASR requerem apenas uma sequência de ácido nucleico de pelo menos 15 nucleotídeos de comprimento. Os ácidos nucleicos que codificam os polipeptídeos ASR podem ser usados como polimorfismo de comprimento do fragmento de marcadores de restrição (RFLP). Southern blots (Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico da planta digerido por restrição pode ser sondado com os ácidos nucleicos que codificam ASR. Os padrões de grupos resultantes podem então ser submetidos à análise genética usando os programas de computador tal como

MapMaker (Lander et al. (1987) Genomics 1: 174-181) a fim de construir um mapa genético. Além disso, os ácidos nucleicos podem ser usados para sondar Southern blots contendo o DNA genômico tratado por endonuclease por restrição de uma série de indivíduos que representam o padrão e progênie de um cruzamento genético definido. A segregação do polimorfismo de DNA é notado e usado para calcular a posição do ácido nucleico que codifica o polipeptídeo ASR no mapa genético previamente obtido usando esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

A produção e uso das sondas derivadas do gene vegetal para o uso no mapeamento genético é descrito em Bernatzky and Tanksley (1986) Plant Mol. Biol. Reporter 4: 37-41. Numerosas publicações descrevem o mapeamento genético dos clones de cDNA específicos usando a metodologia resumida acima ou variações destas. Por exemplo, populações intercruzadas F2, populações cruzadas novamente, populações aleatoriamente auxiliares, linhas isogênicas próximas, e outras séries de indivíduos podem ser usados para o mapeamento. Tais metodologias são bem conhecidas àqueles habilitados na técnica.

As sondas de ácido nucleico também podem ser usadas para o mapeamento físico (isto é, disposição das sequências nos mapas físicos; ver Hoheisel et al. In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346, e referências citadas neste).

Em uma outra forma de realização, as sondas de ácido podem ser usadas na fluorescência direta do mapeamento da hibridização in situ (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora os métodos correntes do mapeamento FISH favorece o uso de amplos clones (diversos kb à diversas centenas de kb; ver Laan et al. (1995) Genome Res. 5:13-20), melhoramento na sensibilidade pode permitir o desempenho do mapeamento FISH usando as sondas mais curtas.

Uma variedade da amplificação do ácido nucleico com base

em métodos para o mapeamento físico e genético pode ser realizado usando o ácidos nucleicos. Exemplos incluem a amplificação específica por alelos (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismo dos fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332), ligação específica por alelos (Landegren et al. (1988) Science 241:1077-1080), reações de extensão por nucleotídeos (Sokolov (1990) Nucleic Acid Res. 18:3671), mapeamento híbrido de radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e Mapeamento Happy (Dear and Cook (1989) Nucleic Acid Res. 17:6795-6807). Para estes métodos, a sequência de um ácido nucleico é usada para projetar e produzir pares de iniciadores para o uso na reação de amplificação ou nas reações da extensão iniciadora. O projeto de tais iniciadores é bem conhecido àqueles habilitados na técnica. Em métodos que utilizam o mapeamento genético com base em PCR, este pode ser necessário para identificar as diferenças da sequência de DNA entre os padrões do cruzamento do mapeamento na região correspondente à sequência do ácido nucleico imediata. Este, entretanto, não é necessário no geral para os métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo traços intensificados relacionados com o rendimento, como descritos anteriormente acima. Estes traços também podem ser combinados com outros traços economicamente vantajosos, tal como traços que itensificam a produção adicional, tolerância a outras tensões bióticas e abióticas, traços que modificam várias características arquiteturais e/ou características bioquímicas e/ou fisiológicas.

VII. Polipeptídeo de fator de transcrição 11 (SPL11) semelhante à proteína promotora Squamosa

Surpreendentemente, tem sido agora observado que a a expressão de modulação em uma planta de um ácido nucleico que codifica um polipeptídeo SPL11 dão plantas tendo os traços relacionados a produção

itensificada relativos às plantas de controle. De acordo com a primeira forma de realização, a presente invenção fornece um método para a intensificação dos traços relacionados de produção em plantas relativos às plantas de controle, que compreende uma expressão de modulação em uma planta de um ácido nucleico que codifica um polipeptídeo SPL11.

A presente invenção também fornece ácidos nucleicos que codificam SPL11 até agora não conhecidos e polipeptídeos SPL11. Estas sequências também sendo úteis na realização de métodos da invenção.

Portanto de acordo com uma forma de realização adicional da presente invenção aqui é fornecido uma molécula de ácido nucleico isolada que compreende:

(i) um ácido nucleico representado pela SEQ ID N°: 448;

(ii) o complemento de um ácido nucleico representado pela SEQ ID N°: 448;

(iii) um ácido nucleico que codifica um polipeptídeo SPL11 tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à sequência de aminoácido representada pela SEQ ID N°: 449 e tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à SEQ ID N°: 465:

SYCQVEGCRTDLSSAKDYHRKHRVCEPHSKAPKVVVA GLERRFCQQCSRFHGLAEFDQKKKSCRRRLNDHNARRRKPQPEAL (que representa o domínio SBP na SEQ ID N°: 449);

um ácido nucleico que hidrodiza sob as condições estringentes à SEQ ID N°: 448.

Além disso, aqui também é fornecido um polipeptídeo isolado que compreende:

(i) uma sequência de aminoácido representada pela SEQ ID

N°: 449;

(ii) uma sequência de aminoácido tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à sequência de aminoácido representada pela SEQ ID N°: 449 e tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à SEQ ID N°: 465:

SYCQVEGCRTDLSSAKDYHRKHRVCEPHSKAPKVVVA GLERRFCQQCSRFHG LAEFDQKKKSCRRRLNDHNARRRKPQPEAL (que representa o domínio SBP na SEQ ID N°: 449).

(iii) derivados de qualquer uma das sequências de aminoácidos dados no (i) ou (ii) acima.

Um método preferido para a expressão modulação (preferivelmente, aumentada) de um ácido nucleico que codifica um polipeptídeo SPL11 é pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo SPL11.

Qualquer referência à seguir a uma "proteína úteis nos métodos da invenção" é deixada significar um polipeptídeo SPL11 como definido neste. Qualquer referência à seguir a um "ácido nucleico úteis nos métodos da invenção" é deixado significar um ácido nucleico capaz de codificar um tal polipeptídeo SPL11. O ácido nucleico será introduzido na planta (e portanto útil na realização de métodos da invenção) é qualquer ácido nucleico que codifica o tipo de proteína que agora será descrita, também neste nomeada "ácido nucleico SPL11" ou "gene SPL11".

Um polipeptídeo SPL11 como definido neste refere-se a um polipeptídeo que compreende um domínio de proteína de ligação Squamosa (SBP), tal domínio tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97 % ou mais identidades de sequências por qualquer um dos domínios SBP como representado pela SEQ

ID N°: 456 a SEQ ID N°: 468 ou SEQ ID N°: 478.

Um "polipeptídeo SPL11" como definido neste compreende a proteína representada pela SEQ ID N°: 448 que equivale a SEQ ID N°: 172 e aos homólogos (ortólogos e paralógos) destes. Preferivelmente, os homólogos da SEQ ID N°: 172 tem um domínio de ligação de DNA. Os polipeptídeos SPL11 podem ser observados em bancos de dados especializados tal como Pfam, (Finn et al. Nucleic Acids Research (2006) Bancos de dados Issue 34:D247-D251). Pfam coleciona uma ampla coleção de alinhamentos de sequência múltipla e modelos Markov secretos (HMM) que cobrem muitos dos domínios de proteína comuns e famílias e é disponível através do Sanger Institute in the United Kingdom.

O corte de ajuntamento limiar do domínio SBP nos modelos Pfam HMM_fs e Pfam HMM_ls é de 25,0. Os pontos curados como considerado nos bancos de dados Pfam são aquelas contagens de sequências mais altas do que o corte de ajuntamento limiar. Entretanto os pontos potenciais, que compreendem os domínios SBP verdadeiros, ainda podem cair sob o corte de ajuntamento. Preferivelmente um polipeptídeo SPL11 é uma proteína tendo um ou mais domínios em sua sequência que excede o corte de ajuntamento da família do domínio de proteína Pfam PF03110, conhecido como domínio SBP.

Alternativamente, um domínio SBP em um polipeptídeo pode ser identificado pela realização de uma comparação de sequência com polipeptídeos conhecidos que compreendem um domínio SBP e estabelecem uma similaridade na região do domínio SBP. As sequências podem ser alinhadas usando qualquer um dos métodos bem conhecidos na técnica tal como algorítmos Blast. A probabilidade para o alinhamento para ocorrer com uma sequência dada é tirado como base para a identificação dos polipeptídeos similares. Um parâmetro que é tipicamente usado ára representar tal probabilidade é denominado valor e. O valor E é uma medida de

confiabilidade da contagem S. A contagem S é uma medida de similaridade em questão à sequência mostrada. O valor E descreve o quanto frequentemente uma contagem dada S é esperado ocorrer aleatoriamente. O corte do valor E pode ser tão alto quanto 1,0. O limiar típico para um bom valor e de uma produção de busca BLAST usando um polipeptídeo SPL11 como sequência em questão pode ser menor do que $e^{-5}(=10^{-5})$ $1.e^{-10}$, $1.e^{-15}$,$1.e^{-20}$, $1.e^{-25}$, $1.e^{-50}$, $1.e^{-75}$, $1.e^{-100}$, $1.e^{-200}$, $1.e^{-300}$, $1.e^{-400}$, $1.e^{-500}$, $1.e^{-600}$, $1.e^{-700}$ e $1.e^{-800}$. Preferivelmente os polipeptídeos SPL11 da invenção compreendem uma sequência tendo uma ordem aumentada de preferência de um valor e inferio do que $e^{-5}(=10^{-5})$, $1.e^{-10}$, $1.e^{-15}$, $1.e^{-20}$, $1.e^{-25}$, $1.e^{-50}$ $1.e^{-75}$, $1.e^{-100}$, $1.e^{-200}$, $1.e^{-300}$, $1.e^{-400}$, $1.e^{-500}$, $1.e^{-600}$, $1.e^{-700}$ e $1.e^{-800}$ em um alinhamento com um domínio SBP observado em um polipeptídeo SPL11 conhecido.

Exemplos de polipeptídeos SPL11 úteis nos métodos da invenção são dados na Tabela G1. As coordenadas do aminoácido do domínio SBP na proteína SPL11 representativa da Tabela G1 são dadas na Tabela G4 é a sequência de domínio é representado pela SEQ ID N°: 456 a SEQ ID N°: 468. Uma sequência de consenso dos domínios SBP presentes nos polipeptídeos SPL11 é dado na SEQ ID N°: 478.

Os polipeptídeos SPL11 preferidos da invenção são aqueles tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, ou mais identidades de sequências por qualquer um dos polipeptídeos dados na Tabela G2.

Tipicamente, os polipeptídeos SPL11 podem compreender em adição ao domínio SBP um ou mais dos seguintes motivos conservados em posições conservadas na sequência relativa ao domínio SBP: (i) Motivo 1 como representado pela SEQ ID N°: 469, (ii) Motivo 2 é uma região rica em serina tipicamente observada na extremidada do terminal N do domínio SBP e pode ser representado pela SEQ ID N°: 470; (iii) Motivo 3 como apresentado pela SEQ ID: 471 que é tipicamente codificada pelos nucleotídeos

compreendidos dentro da região de polipeptídeo SPL11 alvejada pelos membros da família de microRNA miR156; (iv) Motivo 4 como representado pela SEQ ID N°: 472. Figura 27 mostra os motivos conservados e sua posição relativa na sequência de polipeptídeo SPL11 representado pela SEQ ID N°: 428.

Portanto, os polipeptídeos SPL11 preferidos úteis no método da invenção, que compreende em adição ao domínio SBP qualquer um ou mais dos seguintes motivos conservados:

(i) Motivo 1 como representado pela SEQ ID N°: 469 em que qualquer substituição de aminoácido conservativo e/ou 1 ou 2 substituições não conservativas são permitidas,

(ii) Motivo 2 como representado pela SEQ ID N°: 470 em que qualquer substituição de aminoácido é permitida, fornecido que em pelo menos 4 aminoácidos tem uma cadeia secundária polar, preferivelmente serina ou treonina, e fornecido que este motivo é localizado na extremidade do terminal N do domínio SBP;

(iii) Motivo 3 como representado pela SEQ ID: 471, em que 1 ou 2 más combinações são permitidas;

(iv) Motivo 4 como representado pela SEQ ID: 472, em que 1, 2 ou 3 más combinações são permitidas.

Exemplos de substituições de aminoácidos conservativos são dados na seção de fundamento. Tipicamente, os aminoácidos compreendidos em polipeptídeos são aminoácidos alfa tendo uma amina e um grupo carboxila ligado ao mesmo carbono, o carbono alfa, o qual as moléculas de aminoácidos frequentemente compreendem uma cadeia secundária ligada ao carbono alfa. Tabela 3 mostra a classificação dos aminoácidos com base nas propriedades físicas e bioquímicas da cadeia secundária.

Tabela 3. Classificação de aminoácidos de acordo com as propriedades da cadeia secundária.

Aminoácido	3 letras	1 letra	Polaridade da cadeia secundária	Acidez ou basicidade da cadeia secundária	Índice de hidropatia
Arginina	Arg	R	polar	básico	-4,5
Asparagina	Asn	N	polar	neutro	-3,5
Ácido aspártico	Asp	D	polar	ácido	-3,5
Cisteína	Cys	C	polar	neutro	2,5
Ácido glutâmico	Glu	E	polar	ácido	-3,5
Glutamina	Gln	Q	polar	neutro	-3,5
Histidina	His	H	polar	básico	-3,2
Lisina	Lys	K	polar	básico	-3,9
Serina	Ser	S	polar	neutro	-0,8
Treonina	Thr	T	polar	neutro	-0,7
Tirosina	Tyr	Y	polar	neutro	-1,3
Alanina	Ala	A	Não polar	neutro	1,8
Glicina	Gly	G	Não polar	neutro	-0,4
Isoleucina	Ile	I	Não polar	neutro	4,5
Leucina	Leu	L	Não polar	neutro	3,8
Metionina	Met	M	Não polar	neutro	1,9
Fenilalanina	Phe	F	Não polar	neuto	2,8
Prolina	Pro	P	Não polar	neutro	-1,6
Triptofano	Trp	W	Não polar	neutro	-0,9
Valina	Val	V	Não polar	neutro	4,2

Exemplos de polipeptídeos SPL11 que compreendem um ou mais dos motivos conservados do Motivo 1 ao Motivo 4 são dados no Exemplo 62. Figura 28 mostra a posição dos motivos conservados naqueles polipeptídeos SPL11.

Preferivelmente, o polipeptídeo SPL11 da invenção quando usado na construção de uma árvore filogenética, tal como um descrito na Figura 29, os grupos dentro do grupo S3 que compreende a sequência de aminoácido representada pela SEQ ID N°: 428 (nomeada AtSPL11 na Figura 29) antes do que com qualquer outro grupo.

O termo "domínio" e "motivo" é definido nas seções de "definições" neste. Os bancos de dados especializados existem para a identificação dos domínios, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-244, InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318,

Prosite (Bucher and Bairoch (1994), Uma sintaxe de perfil generalizada for biomolecular sequences motivos e sua função em interpretação de sequência automática. (In) ISMB-94; Proceedings 2° International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D. eds., pp53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004), ou Pfam (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002). Uma série de ferramentaspara análise *in silico* das sequências de proteínas é disponível no servidor proteômico ExPASy (Swiss Institute of Bioinformatics (Gasteiger et al. exPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31:3784-3788(2003)). Os domínios também podem ser identificados usando as técnicas de rotina, tal como pelo alinhamento da sequência.

Os métodos para o alinhamento das sequências em comparação são bem conhecidos na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. O uso de GAP do algorítmo de Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) observam o alinhamento global (isto é medido-se as sequências completas) de duas sequências que maximizam o número de pontos e minimizam o número de fendas. O algorítmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-10) calcula o percentual da identidade de sequência e realiza uma análise estatística da similaridade entre as duas sequências. O software para a realização da análise de BLAST é publicamente disponível através de National Centre for Biotechnology Information (NCBI). Os homólogos podem ser prontamente identificados usando, por exemplo, o algorítmo de alinhamento de sequência múltipla ClustalW (versão 1.83), com os parâmetros de alinhamento em pares default, er um método de contagem em porcentagem. As porcentagens globais de similaridade e identidade também podem ser determinadas usando um dos métodos disponíveis na embalagem do software MatGAT (Campanella et al., BMC Bioinformatics. 2003 Jul

10;4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando protein or DNA sequences). Uma edição menor do manual pode ser realizada para otimizar o alinhamento entre motivos conservados, como deve ser evidente a uma pessoa habilitada na técnica. Além disso, em vez do uso das sequências de comprimento total para a identificação de homólogos, domínios específicos também podem ser usados. Os valores da identidade de sequência podem ser determinados no a´cido nucleico inteiro ou sequência de aminoácido ou em domínios selecionados ou motivos conservados, usando os programas mencionados acima usando os parâmetros default.

Além disso, os polipeptídeos SPL11 (em pelo menos na sua forma natural) pode ter tipicamente a atividade de ligação de DNA, em particular estes podem ligar os fragmentos de DNA que compreendem a caixa de ligação do DNA do domínio SBP como representado pela SEQ ID N°: 49. Tipicamente a caixa de ligação do DNA do domínio SBP é observada nos promotores de genes de origem vegetal tal como o gene SQUAMOSA. Os fragmentos que compreendem a caixa de ligação do DNA do domínio SBP são preferivelmente mais do que 10, 15, 20, 25, 100, 200, 500, 1000, 2000 longos pares de bases. Os métodos para determinar a ligação de DNA do domínio SBP contendo as proteínas são aplicáveis aos polipeptídeos SPL11 e são conhecidos na técnica (Klein at al. Mol Gen Genet. 1996, 15;250(1):7-16; Yamasaki et al. 2004 J Mol Biol. 2004, 12;337(1):49-63).

Os polipeptídeos SPL11 preferidos da invenção são aqueles tendo a atividade de ligação de DNA, mais preferivelmente aquela ligação de um fragmento de DNA que compreende a SEQ ID N°: 475.

A presente invenção é ilustrada pela transformação das plantas com a sequência do ácido nucleico representada pela SEQ ID N°: 427, que codifica a sequência de polipeptídeo da SEQ ID N°: 428. Entretanto, o desempenho da invenção não é restrito a estas sequências; os métodos da

invenção podem ser vantajosamente realizados usando qualquer SPL11 que codifica o ácido nucleico ou polipeptídeo SPL11 como definido neste.

Exemplos de ácidos nucleicos que codifica os polipeptídeos SPL11 são dados na Tabela G1 do Exemplo 62 neste. Tais ácidos nucleicos são úteis na realização de métodos da invenção. A sequência de aminoácidos dados na Tabela G1 do Exemplo 62 são as sequências exemplo de ortólogos e paralógos do polipeptídeo SPL11 representado pela SEQ ID N°: 428, os termos "ortólogos" e "parálogos" sendo como definido neste. Ainda os ortólogos e paralógos podem ser prontamente identificados pela realização de uma busca blast recíproca denominada. Tipicamente, este envolve um primeiro BLAST que envolve uma sequência em questão BLASTing (por exemplo usando qualquer uma das sequências listadas na Tabela G1 do Exemplo 62) contra qualquer base de dados da sequência, tal como publicamente disponível pelos bancos de dados NCBI. BLASTN ou TBLASTX (usando valores default padrões) são geralmente usados quando começa-se a partir de uma sequência de nucleotídeo, e BLASTP ou TBLASTN (usando valores default padrões) quando começa-se a partir de uma sequência de proteína. Os resultados BLAST podem ser opcionalmente filtrados. As sequências de comprimento total dos resultados filtrados ou resultados não filtrados são então novamente submetidos ao BLAST (segundo BLAST) contra a sequência de organismo de que a sequência em questão é derivada (onde a sequência em questão é a SEQ ID N°: 427 ou SEQ ID N°: 428, o segundo BLAST portanto seria contra as sequências de Arabidopsis). Os resultados dos primeiros e segundos BLAST são então comparados. Um parálogo é identificado se um acerto de classificação alto a partir do primeiro blast é as mesmas espécies como de que a sequência em questão é derivada, um BLAST, novamente, então resultam de maneira ideal na sequência em questão entre os acertos mais altas; um ortólogo é identificado se um acerto de classificação alto no primeiro BLAST não é das mesmas espécies como de

que a sequência em questão é derivada, e preferivelmente resultam no BLAST novamente na sequência em questão sendo entre os acertos mais altos.

Os acertos de classificações altos são aqueles tendo um valor E baixo. O valor E inferior, a contagem mais significante (ou em outras palavras a possibilidade inferior que o acerto foi observado pela possibilidade). A computação do valor e é bem conhecido na técnica. Além disso aos valores E, em comparação são também contados pela identidade de porcentagem. A identidade da porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácido nucleico comparado (ou polipeptídeo) em um comprimento particular. No caso de amplas famílias, ClustalW pode ser usado, seguido por uma árvore de ligação próxima, para ajudar a visualizar o agrupamento dos genes relacionados e para identificar ortólogos e paralógos.

As variantes do ácido nucleico também pode ser útil na prática dos métodos da invenção. Exemplos de tais variantes incluem ácidos nucleicos que codificam os homólogos e derivados de qualquer uma da sequência de aminoácidos dados na Tabela G1 do Exemplo 62, os termos "homólogo" e "derivado" sendo como definido neste. Também úteis nos métodos da invenção são os ácidos nucleicos que codificam os homólogos e derivados de ortólogos ou parálogos de qualquer uma da sequência de aminoácidos dados na Tabela G1 do Exemplo 62. Homólogos e derivados úteis nos métodos da presente invenção tem substancialmente a mesma atividade funcional e biológica como a proteína não modificada de que estes são derivados.

Ainda as variantes do ácido nucleico útil na prática dos métodos da invenção incluem as porções dos ácidos nucleicos que codificam os polipeptídeos SPL11, ácidos nucleicos que hibridizam aos ácidos nucleicos que codificam os polipeptídeos SPL11, variantes de união de ácidos nucleicos que codificam os polipeptídeos SPL11, variantes alélicas de ácidos nucleicos

que codificam os polipeptídeos SPL11 e variantes de ácidos nucleicos que codificam os polipeptídeos SPL11 obtido pela mistura do gene. Os termos que hidridiza a sequência, variante de união, variante alélica e mistura do gene são como descritos neste.

Uma variante de ácido nucleico preferido útil na prática dos métodos da invenção é um ácido nucleico que codifica um polipeptídeo SPL11, que é microRNA insensível, ainda preferivelmente o ácido nucleico é insensível às propriedades de microRNAs à família miR156. Ácidos nucleicos alvos MicroRNA (RNA em particular) para a destruição tipicamente causando uma redução ou inibição do acúmulo do RNA alvejado. O alvejamento requer a hibridização entre o microRNA e o ácido nucleico alvejado (RNA) em um local alvo da região muito específica, denominado o miR microRNA), que compreende uma sequência complementar a uma porção do gene microRNA maduro. Tipicamente os ácidos nucleicos microRNA insensíveis compreendem uma sequência tendo uma ordem aumentada de preferência 1, 2, 3, 4, 5 ou mais más combinações em um alinhamento à molécula microRNA relevante no local alvo miR. Os ácidos nucleicos insensíveis ao MicroRNA podem acumular em uma célula aos níveis em uma ordem aumentada de preferência 5, 10, 20, 30, 40, 50 vezes ou mais altas do que o ácido nucleico correspondente alvejado pelo MicroRNA relevante, que tipicamente compreende 100 % da sequência complementar no local alvo miR.

A família miR156 foi descrita precoce e uma compilação dos microRNA incluindo uma família de membros miR156 pode ser observado em bancos de dados miRBase (Griffiths-Jones et al. 2006 Nucleic Acids Research, 2006, Vol. 34, Database issue D140-D144). Os bancos de dados miRBase é mantido pelo The Wellcome Trust Sanger Institute, in Cambridge, UK. O local alvo miR156 em um ácido nucleico SPL11 é complementar à sequência madura de miR156 microRNAs. Um exemplo das sequências

maduras de uma família de membros miR156 de arroz e seus locais alvos correspondentes em ácidos nucleicos SPL de arroz é dado na Figura 30 A. Figura 31 B mostra um alinhamento dos ácidos nucleicos SPL11 representativos (na forma de DNA), em que o local alvo de miR156 nos riboácidos nucleicos correspondentes é indicado. Um exemplo de um ácido nucleico SPL11 insensível miR156 que codifica a SEQ ID N°: 428 é representado pela SEQ ID N°: 431. Ainda os exemplos de ácido nucleico SPL11 insensível miR156s são representados pela SEQ ID 440 e SEQ ID N°: 454, o último necessitado do local alvo miR156.

Preferivelmente um ácido nucleico SPL11 útil nos métodos da invenção tem 1, 2, 3, 4 ou mais más combinações no local alvo miR156 ou perda do local alvo do gene miR156. Exemplos de tais ácidos nucleicos SPL11 são dados na Figura 31 B. Ainda preferivelmente é um ácido nucleico como representado pela SEQ ID N°: 431, SEQ ID 440 e SEQ ID N°: 454.

Os ácidos nucleicos que codificam os polipeptídeos SPL11 necessitam ser os ácidos nucleicos de comprimento total, visto que o desempenho dos métodos da invenção não confiam no uso das sequências de ácido nucleico de comprimento total. De acordo com uma presente invenção, aqui é fornecido um método para a intensificação dos traços relacionados de produção em plantas, que compreende a introdução e expressão em uma planta de uma porção de qualquer uma da sequência dos ácidos nucleicos dados na Tabela G1 do Exemplo 62, ou uma porção de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma da sequência de aminoácidos dados na Tabela G1 do Exemplo 62.

Uma porção de um ácido nucleico pode ser preparada, por exemplo, pela fabricação de um ou mais anulações ao ácido nucleico. As porções podem ser usadas na forma isolada ou este pode ser fundida a outras sequências codificadoras (ou não codificadoras) a fim de, por exemplo, produzir uma proteína que combina de diversas atividades. Quando fundido a

outras sequências codificadoras, o polipeptídeo resultante produzido na produção pode ser maior do que predito para a porção de proteína.

As porções úteis nos métodos da invenção codificam um polipeptídeo SPL11 como definido neste, e tem substancialmente a mesma atividade biológica como uma sequência de aminoácidos dados na Tabela G1 do Exemplo 62. Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos dados na Tabela G1 do Exemplo 62, ou é uma porção de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma da sequência de aminoácidos dados na Tabela G1 do Exemplo 62. Preferivelmente a porção é em pelo menos 70, 100, 200, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000 de nucleotídeos consecutivos em comprimento, os nucleotídeos consecutivos sendo de qualquer uma da sequência dos ácidos nucleicos dados na Tabela G1 do Exemplo 62, ou de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma da sequência de aminoácidos dados na Tabela G1 do Exemplo 62. Preferivelmente a porção codifica em pelo menos um domínio SBP. Mais preferivelmente a porção é uma porção do ácido nucleico da SEQ ID N°: 427. Preferivelmente, a porção codifica uma sequência de aminoácido o qual quando usado na construção de uma árvore filogenética, tal como um descrito na Figura 29, os grupos com o grupo de polipeptídeos SPL11 que compreendem a sequência de aminoácido representada pela SEQ ID N°: 428 (AtSPL11) antes do que com qualquer outro grupo.

Uma outra variante de ácido nucleico útil nos métodos da invenção é um ácido nucleico capaz de hibridizar, sob condições de estringência reduzida, preferivelmente sob condições estringentes, com um ácido nucleico que codifica um polipeptídeo SPL11 como definido neste, ou com uma porção como definido neste.

De acordo com uma presente invenção, aqui é fornecido um

método para a intensificação dos traços relacionados de produção em plantas, que compreende a introdução e expressão em uma planta de um ácido nucleico capaz da hidridização a qualquer um dos ácidos nucleicos dados na Tabela G1 do Exemplo 62, ou que compreende a introdução e expressão em uma planta de um ácido nucleico capaz de hibridizar a um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma da sequência dos ácidos nucleicos dados na Tabela G1 do Exemplo 62.

A hibridização da sequências útil nos métodos da invenção codificam um polipeptídeo SPL11 como definido neste, e tem substancialmente a mesma atividade biológica como uma sequência de aminoácidos dados na Tabela G1 do Exemplo 62. Preferivelmente, a hidridização da sequência é capaz de hibridizar a qualquer um dos ácidos nucleicos dados na Tabela G1 do Exemplo 62, ou a uma porção de qualquer uma destas sequências, uma porção sendo como definido acima, ou em que a hidridização da sequência é capaz de hibridizar a um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma da sequência de aminoácidos dados na Tabela G1 do Exemplo 62. Mais preferivelmente, a hidridização da sequência é capaz de hibridizar a um ácido nucleico como representado pela SEQ ID N°: 427 ou a uma porção destes.

Preferivelmente, a hidridização da sequência codifica uma sequência de aminoácido o qual quando usado na construção de uma árvore filogenética, tal como um descrito na Figura 29, os grupos com o grupo de polipeptídeos SPL11 que compreende a sequência de aminoácido representada pela SEQ ID N°: 428 (AtSPL11) antes do que com qualquer outro grupo.

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante de união que codifica um polipeptídeo SPL11 como definido anteriormente acima, a variante de união sendo como definido neste.

De acordo com uma presente invenção, aqui é fornecido um

método para a intensificação dos traços relacionados de produção em plantas, que compreende a introdução e expressão em uma planta de uma variante de união de qualquer uma da sequência dos ácidos nucleicos dados na Tabela G1 do Exemplo 62, ou a variante de união de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma da sequência de aminoácidos dados na Tabela G1 do Exemplo 62. Exemplos de variantes de união do gene que codifica a SEQ ID N°: 428 são representados pela SEQ ID N°: 427; SEQ ID N°: 429 e SEQ ID N°: 430.

As variantes de união preferidas são as variantes de união de um ácido nucleico representado pela SEQ ID N°: 427, ou a variante de união de um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 428. Preferivelmente, a sequência de aminoácido codificada pela variante de união, quando usado na construção de uma árvore filogenética, tal como um descrito na Figura 29, os grupos com o grupo de polipeptídeos SPL11 que compreende a sequência de aminoácido representada pela SEQ ID N°: 428 (AtSPL11) antes do que com qualquer outro grupo.

Uma outra variante de ácido nucleico útil na realização de métodos da invenção é uma variante alélica de um ácido nucleico que codifica um polipeptídeo SPL11 como definido anteriormente acima, de uma variante alélica sendo como definido neste.

De acordo com uma presente invenção, aqui é fornecido um método para a intensificação dos traços relacionados de produção em plantas, que compreende a introdução e expressão em uma planta de uma variante alélica de qualquer um dos ácidos nucleicos dados na Tabela G1 do Exemplo 62, ou que compreende a introdução e expressão em uma planta de uma variante alélica de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma da sequência de aminoácidos dados na Tabela G1 do Exemplo 62.

As variantes alélicas úteis nos métodos da presente invenção

tem substancialmente a mesma atividade biológica como o polipeptídeo SPL11 da SEQ ID N°: 428 e qualquer um dos aminoácidos descrito na Tabela G1 do Exemplo 62. As variantes alélicas existem na natureza, e abrage dentro dos métodos da presente invenção é o uso destes alelos naturais. Preferivelmente, a variante alélica é uma variante alélica da SEQ ID N°: 427 ou uma variante alélica de um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N°: 428. Preferivelmente, a sequência de aminoácido codificada pela variante alélica, quando usada na construção de uma árvore filogenética, tal como um descrito na Figura 29, os grupos com os polipeptídeos SPL11 que compreendem a sequência de aminoácido representada pela SEQ ID N°: 428 (AtSPL11) antes do que com qualquer outro grupo.

A mistura do gene ou evolução direcionada também pode ser usada para gerar as variantes de ácidos nucleicos que codificam os polipeptídeos SPL11 como definido acima; o termo "a mistura do gene" sendo como definido neste.

De acordo com uma presente invenção, aqui é fornecido um método para a intensificação dos traços relacionados de produção em plantas, que compreende a introdução e expressão em uma planta de uma variante de qualquer uma das sequências de ácidos nucleicos dados na Tabela G1 do Exemplo 62, ou que compreende a introdução e expressão em uma planta de uma variante de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma da sequência de aminoácidos dados na Tabela G1 do Exemplo 62, no qual da variante do ácido nucleico é obtido pela mistura do gene.

Preferivelmente, a sequência de aminoácido codificada pela variante de ácido nucleico obtido pela mistura do gene, quando usado na construção de uma árvore filogenética tal como um descrito na Figura 29, os grupos com o grupo de polipeptídeos SPL11 que compreende a sequência de

aminoácido representada pela SEQ ID N°: 428 (AtSPL11) antes do que com qualquer outro grupo.

Além disso, as variantes do ácido nucleico também podem ser obtidos pela mutagênese direcionada ao local. Diversos métodos são disponíveis para atingir a mutagênese direcionada ao local, a mais comum sendo métodos com base em PCR (Current Protocols in Molecular Biology. Wiley Eds.).

Os ácidos nucleicos que codificam os polipeptídeos SPL11 podem ser derivados de qualquer fonte natural ou artificial. O ácido nucleico pode ser modificada a partir da forma natural na composição e/ou ambiente genômico através da consideração da manipulação humana. Preferivelmente o ácido nucleico que codifica o polipeptídeo SPL11 é a partir de uma planta, ainda preferivelmente de uma planta dicotiledônia, mais preferivelmente da família Brassicaceae, mais preferivelmente o ácido nucleico é de *Arabidopsis thaliana.*

O desempenho dos métodos da invenção dão plantas tendo os traços relacionados a produção itensificada. Em particular o desempenho dos métodos da invenção dão plantas tendo a produção aumentada especialmente a produção da semente aumentada relativa às plantas de controle. Os termos "produção" e "produção de semente" são descritos em mais detalhes na seção "definições" neste.

A referência neste à traços relacionados a produção itensificada é deixada significar um aumento na biomassa (peso) de um ou mais partes de uma planta, que pode incluir partes acima do solo (coletáveis) e/ou partes abaixo do solo (coletáveis). Em particular, tais partes colhíveis são semeadas, e o desempenho dos métodos da invenção resultam em plantas tendo a produção aumentada relativa à produção de plantas de controle.

A retirada de milho como um exemplo, uma produção aumentada pode ser manifestada como um ou mais dos seguintes: aumento no

número de plantas estabelecido por hectare ou acre, um aumento no número de espigas por plantas, um aumento no número de séries, number de núcleos por célula, peso do núcleo, comprimento/diâmetro do núcleo da espiga de mil pesos, aumento na taxa de enchimento da semente (que é o número de sementes enchidas dividas pelo número total de sementes e multiplicadas por 100), entre outros. A retirada do arroz como um exemplo, uma produção aumentada pode manifestar por si própria como um aumento em um ou mais dos seguintes: número de plantas por hectare ou acre, número de panículas por planta, número de espigas por panículas, número de flores (florzinhas) por panícula (que é expressado como uma razão do número de sementes enchidas no número de panículas primárias), aumento na taxa de enchimento da semente (que é o número de sementes enchidas dividas pelo número total de sementes e multiplicadas por 100), aumento em peso de núcleo milhar, entre outros.

A presente invenção fornece um método para a produção aumentada especialmente a produção de semente de plantas, relativa às plantas de controle, que o método compreende uma expressão de modulação, preferivelmente a expressão aumentada em uma planta de um ácido nucleico que codifica um polipeptídeo SPL11 como definido neste.

Visto que as plantas transgênicas de acordo com uma presente invenção tem a produção aumentada, é igualmente que estas plantas exibem uma taxa de desenvolvimento aumentada (durante pelo menos parte de sei cuclo de vida), relativo à taxa de desenvolvimento de plantas de controle em um estágio correspondente em seu ciclo de vida.

A taxa de desenvolvimento aumentada pode ser específica por uma ou mais partes de uma planta (incluindo sementes), ou pode ser por toda parte substancialmente do total da planta. As plantas tendo uma taxa de desenvolvimento aumentada podem ter um ciclo de vida mais curto. O ciclo de vida de uma planta pode ser levado significar o período necessário para o

desenvolvimento a partir da semente madura seca até o estágio onde a planta tem produzido sementes maduras secas, similar ao material de partida. Este ciclo de vida pode ser influenciado pelos fatores tal como vigor precoce, taxa de desenvolvimento, índice de verdor, período e velocidade de florescência do amadurecimento da semente. O aumento na taxa de desenvolvimento pode acontecer em um ou mais estágios no ciclo de vida de uma planta ou durante substancialmente o ciclo de vida total da planta. A taxa de desenvolvimento aumentada durantes os estágios precocemente no ciclo de vida de uma planta pode refletir o vigor itensificado. O aumento na taxa de desenvolvimento pode alterar o ciclo de coleta de uma planta permitindo que as plantas sejam semeadas por último e/ou coletadas mais cedo do que seria possível de outra maneira (um efeito similar pode ser obtido com o período de florescência precoce). Se a taxa de desenvolvimento é suficientemente aumentada, pode permitir a semeadura adicional das sementes das mesmas espécies de plantas (por exemplo semeadura e coletagem de plantas de arroz seguido pela semeadura e coletagem de plantas de arroz adicionais todas dentro de um período de desenvolvimento convencional). Similarmente, se a taxa de desenvolvimento é suficientemente aumentada, pode permitir a semeadura adicional das sementes de espécies de plantas diferentes (por exemplo a semeadura e coletagem de plantas de milho seguidas, por exemplo, pela semeadura e coletagem opcional de soja, batata ou qualquer outra planta adequada). Os períodos adicionais de coletagem a partir do mesmo rizoma no caso de algumas de plantas de cultivo também podem ser possíveis. A alteração do ciclo de coleta de uma planta pode levar a um aumento na produção de biomassa anual por acre (devido a um aumento no número de vezes (dito em um ano) que qualquer planta particular pode ser desenvolvida e coletada). Um aumento na taxa de desenvolvimento também pode permitir para o cultivo de plantas transgênicas em uma área geográfica extensa do que sua contra-parte de tipo selvagem, visto que as limitações territoriais para o

desenvolvimento de uma lavoura são frequentemente determinados pelas condições ambientais adversas no período do plantio (temporada antecipada) ou no período de coleta (temporada tardia). Tais condições adversas podem ser evitadas se o ciclo coletado é encurtado. A taxa de desenvolvimento pode ser determinado pela derivação de vários parâmetros a partir das curvas de desenvolvimento, tais parâmetros podem ser: T-Mid (o período de retirada para as plantas para atingir 50 % de seu tamanho máximo) e T-90 (período de retirada para as plantas para atingir 90 % de seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção dão plantas tendo uma taxa de desenvolvimento aumentada relativas às plantas de controle. Portanto, de acordo com uma presente invenção, aqui é fornecido um método para a taxa de desenvolvimento aumentada das plantas, que o método compreende uma expressão de modulação, preferivelmente a expressão aumentada em uma planta de um ácido nucleico que codifica um polipeptídeo SPL11 como definido neste.

Um aumento na produção e/ou taxa de desenvolvimento ocorre se a planta está sob as condições de não tensão ou se a planta é exposta à várias tensões em comparação as plantas de controle. As plantas tipicamente respondem a exposição da tensão pelo desenvolvimento mais lentamente. Em condição de diversa tensão, a planta ainda pode interromper completamente o desenvolvimento. A branda tensão em outras palavras é definida neste como sendo qualquer tensão pelo que uma planta é exposta que não resulta no interromper da planta para desenvolver completamente sem a capacidade para resumir o desenvolvimento. A branda tensão no sentido da invenção leva a uma redução no desenvolvimento das plantas submetidas à tensão de menos do que 40 %, 35 % ou 30 %, preferivelmente menos do que 25 %, 20 % ou 15 %, mais preferivelmente menos do que 14 %, 13 %, 12 %, 11 % ou 10 % ou

menos em comparação à planta de controle sob as condições de não tensão. Devido aos avanços nas práticas da agricultura (irrigação, fertilização, tratamentos com pesticidas) diversas tensões não são frequentemente encontradas nas plantas de controle cultivadas. Como uma consequência, o desenvolvimento compromissado induzido pela branda tensão é frequentemente uma característica indesejável para a agricultura. As tensões brandas são as tensões (mabientais) bióticas e/ou abióticas todos os dias pelo qual uma planta é exposta. As tensões abióticas podem ser devido a aridez ou excesso de água, tensão anaeróbica, tensão ao sal, toxicidade química, tensão oxidativa e temperaturas quentes, frias ou de congelamento. A tensão abiótica pode ser uma tensão osmótica causada por uma tensão à água (particularmente devido a aridez), tensão ao sal, tensão oxidativa ou uma tensão iônica. As tensões bióticas são tipicamente aquelas tensões causadas por patogênios, tal como bactéria, vírus, fungos e insetos.

Em particular, os métodos da presente invenção podem ser realizados sob as condições de não tensão ou sob as condições de branda aridez para dar as plantas tendo a produção aumentada relativa às plantas de controle. Como relatado em Wang et al. (Plant (2003) 218: 1-14), as tensões abióticas levam a uma série de mudanças moleculares, bioquímicas, fisiológicas ou morfológicas que adversamente afetam o desenvolvimento e produtividade da planta. Aridez, temperaturas de salinidade extremas e tensão oxidativa são conhecidos serem interconectados o podem induzir o desenvolvimento e dano celular através dos mecanismos similiares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve um alto grau particularmente de "cruzamento" entre tensão à aridez e tensão de alta salinidade. Por exemplo, aridez e/ou salinização são manifestados principalmente como tensão osmótica, resultando no rompimento de homeostase e distribuição de íon na célula. A tensão oxidativa, que frequentemente acompanha a alta e baixa temperatura, salinidade ou tensão à

aridez, pode causar desnaturação de proteínas estruturais e funcionais. Como uma consequência, estas diversas tensões ambientais frequentemente ativam os caminhos de sinalização celular similares e respostas celulares, tal como a produção de proteínas de tensão, super regulação de anti oxidantes, acúmulo de solutos compatíveis e impede o desenvolvimento. O termo condições de "não tensão" como usado neste são aquelas condições ambientais que permitem o ótimo desenvolvimento das plantas. As pessoas habilitadas na técnica são atentas as condições normais do solo e condições climáticas para uma dada localização. As plantas com ótima condições de desenvolvimento, (desenvolvimento sob as condições de não tensão) tipicamente produzido em uma ordem aumentada de preferência em pelo menos 90 %, 87 %, 85 %, 83 %, 80 %, 77 % ou 75 % da produção média de tal planta em um dado ambiente. A produção média pode ser calculada na base de coleta e/ou temporada. As pessoas habilitadas na técnica são atentas as porduções de rendimento médios de uma lavoura.

O desempenho dos métodos da invenção dão as plantas o desenvolvimento sob as condições de não tensão ou sob a produção das condições de branda aridez aumentada relativa às plantas de controle desenvolvidas sob as condições comparáveis. Portanto, de acordo com uma presente invenção, aqui é fornecido um método para a produção aumentada no desenvolvimento das plantas sob as condições de não tensão ou sob as condições de branda aridez, que o método compreende uma expressão aumentada em uma planta de um ácido nucleico que codifica um polipeptídeo SPL11.

O desempenho dos métodos da invenção dão o desenvolvimento das plantas sob condições de deficiência de nutriente, particularmente sob condições de deficiência de nitrogênio, produção aumentada relativa às plantas de controle desenvolvidas sob as condições comparáveis. Portanto, de acordo com uma presente invenção, aqui é

fornecido um método para a produção aumentada no desenvolvimento das plantas sob condições de deficiência de nutriente, que o método compreende uma expressão aumentada em uma planta de um ácido nucleico que codifica um polipeptídeo SPL11. A deficiência de nutriente pode resultar a partir de uma perda ou excesso de nutrientes tal como nitrogênio, fosfatos e outros compostos contendo fósforos, potássio, cálcio, cádmio, magnésio, manganês, ferro e boro, entre outros.

A presente invenção abrange plantas ou partes destas (incluindo sementes) obtidas pelos métodos de acordo com uma presente invenção. As plantas ou partes destas compreendem um transgene de um ácido nucleico que codifica um polipeptídeo SPL11 como definido acima.

A invenção também fornece as construções genéticas e vetores para facilitar a introdução e/ou expressão em plantas de ácidos nucleicos que codificam os polipeptídeos SPL11. As construções dos genes podem ser inseridos em vetores, que podem ser comercialmente disponíveis, adequados para a transformação em plantas e adequados para a expressão do gene de interesse nas células transformadas. A invenção também fornece o uso de uma construção do gene como definido neste nos métodos da invenção.

Mais especificamente, a presente invenção fornece uma construção que compreende:

(a) um ácido nucleico que codifica um polipeptídeo SPL11 como definido acima;

(b) uma ou mais sequências de controle capazes da expressão de condução da sequência do ácido nucleico de (a); e opcionalmente

(c) uma sequência de terminal de transcrição.

Preferivelmente, o ácido nucleico que codifica um polipeptídeo SPL11 é como definido acima. O termo "sequência de controle" e "sequência de terminação" são como definidos neste.

As plantas são transformadas com um vetor que compreende

qualquer um dos ácidos nucleicos descritos acima. O técnico habilitado está bem atento dos elementos genéticos que devem estar presentes no vetor a fim de transformar sucessivamente, selecionar e propagar as células hospedeiras contendo a sequência de interesse. A sequência de interesse é operavelmente ligada a uma ou mais sequências de controle (em pelo menos a um promotor).

Vantajosamente, qualquer tipo de promotor, se natural ou sintético, pode ser usado para conduzir a expressão da sequência do ácido nucleico. Um promotor constitutivo é particularmente útil nos métodos. Ver a seção de "Definições" neste para as definições de vários tipos de promotores.

Será claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico que codifica o polipeptídeo SPL11 representado pela SEQ ID N°: 427, ou é a aplicabilidade da invenção restrita a expressão de um ácido nucleico que codifica o polipeptídeo SPL11 quando conduzido por um promotor constitutivo.

O promotor constitutivo é preferivelmente um promotor GOS2, preferivelmente um promotor GOS2 de arroz. Ainda preferivelmente o promotor constitutivo é representado por uma sequência de ácido nucleico substancialmente similar a SEQ ID N°: 476, mais preferivelmente o promotor constitutivo é como representado pela SEQ ID N°: 476. Ver a seção de "Definições" neste ainda pelos exemplos de promotores constitutivos.

Em uma forma de realização alternativa, um promotor específico de semente é usado. O promotor específico de semente é preferivelmente ABA (ácido abcísico) induzível, é preferivelmente o promotor WSI18, preferivelmente o WSI18 de arroz. Ainda preferivelmente o promotor constitutivo é representado por uma sequência de ácido nucleico substancialmente similar a SEQ ID N°: 477. Ver a seção de "Definições" neste ainda pelos exemplos de promotores específicos de sementes.

Deve ser claro que os promotores úteis nos métodos da invenção não são limitados aqueles especificados nas formas de realização

anteriormente mencionadas.

Opcionalmente, uma ou mais sequências terminadoras podem ser usadas na construção introduzida em uma planta. Os elementos reguladores adicionais podem incluir transcripcionais tão bem quanto itensificadores de tradução. Aqueles habilitados na técnica serão atentos as sequências itensificadoras e terminadoras que podem ser adequadas para o uso na realização da invenção. Uma sequência íntron também pode ser adicionada à região não traduzida 5' (UTR) ou na sequência codificadora para aumentar a quantidade da mensagem madura que acumula no citosol, como descrito na seção de definições. Outra sequência de controle (além do itensificador promotor, silenciador, sequências de íntron, regiões 3'UTR e/ou 5'UTR) pode ser elementos que estabilizam o RNA e/ou proteína. Tais sequências devem ser conhecidas ou podem ser prontamente obtidas por uma pessoa habilitada na técnica.

As construções genéticas da invenção ainda podem incluir uma origem da sequência de replicação que é requerida para a manutenção e/ou replicação em um tipo de célula específica. Um exemplo é quando uma construção genética é requerida ser mantida em uma célula bacteriana como um elemento genético episomal (por exemplo molécula de plasmídeo ou cosmídeo). As origens preferidas de replicação incluem, mas não são limitadas a, a f1-ori e colE1.

Para a detecção da transferência sucessível da sequência dos ácidos nucleicos como usado nos métodos da invenção e/ou seleção de plantas transgênicas que compreendem estes ácidos nucleicos, é vantajoso o uso de genes marcadores (ou genes repórteres). Portanto, a construção genética pode opcionalmente compreender um gene marcador selecionável. Os marcadores selecionáveis são descritos em mais detalhes na seção de "definições" neste.

É conhecido que na integração transiente ou estável de ácidos

nucleicos em células de planta, apenas uma minoria de células absorve o DNA estranho e, se desejado, integra-se a este genoma, dependendo vetor de expressão usado e a técnica de transfecção usada. Para identificar e selecionar estes integrantes, um gene codificador para um marcador selecionável (tal como um descrito acima) é usualmente introduzido nas células hospedeiras junto com o gene de interesse. Estes marcadores podem ser por exemplo usados nos mutantes em que estes genes não são funcionais por, por exemplo, anulação por métodos convencionais. Além disso, as moléculas de ácido nucleico que codificam um marcador selecionável podem ser introduzidas em uma célula hospedeira no mesmo vetor que compreende a sequência que codifica os polipeptídeos da invenção ou usados nos métodos da invenção, ou além disso em um vetor separador. As células que foram transfectadas estáveis com o ácido nucleico introduzido podem ser identificadas por exemplo pela seleção (por exemplo, as células que tem integrado os marcadores selecionáveis sobrevivem considerando as outras células mortas). Os genes marcadores podem ser removidos ou taxados de células transgênicas uma vez que estas não são mais necessárias. As técnicas para a remoção do gene marcador são conhecidas na técnica, as técnicas úteis são descritas acima na seção de definições.

A invenção também fornece um método para a produção de plantas transgênicas tendo os traços relacionados a produção itensificada relativas às plantas de controle, que compreende a introdução e expressão em uma planta de qualquer ácido nucleico que codifica um polipeptídeo SPL11 como definido anteriormente acima.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo os traços aumentados relacionados a produção itensificada, particularmente produção itensificada (semente), que o método compreende:

(i) a introdução e expressão em uma planta ou célula de planta

de um ácido nucleico que codifica o polipeptídeo SPL11; e

(ii) cultivo da célula de planta sob condições que promovem o o desenvolvimento e crescimento vegetal.

O ácido nucleico de (i) pode ser qualquer um dos ácidos nucleicos capazes que codifica um polipeptídeo SPL11 como definido neste.

O ácido nucleico pode ser introduzido diretamente em uma célula de planta ou na planta por si própria (incluindo a introdução em um tecido, orgão ou qualquer outra parte de uma planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é preferivelmente introduzido em uma planta pela transformação. O termo "transformação" é descrito em mais detalhes na seção "definições" neste.

As células de planta modificadas geneticamente podem ser regeneradas por intermédio de todos os métodos com que os trabalhadores habilitados sejam familiares. Os métodos adequados podem ser observados nas publicações anteriormente mencionadas por S.D. Kung and R. Wu, Potrykus ou Hofgen and Willmitzer.

Geralmente após a transformação, células de plantas ou disposições celulares são selecionados pela presença de um ou mais marcadores que são codificados pelos genes expressíveis às plantas co-transferidos com o gene de interesse, seguido que o material transformado é regenerado em um total da planta. Para selecionar as plantas transformadas, o material vegetal obtido na transformação é, como uma regra, submetido às condições seletivas de modo que as plantas transformadas podem ser distinguidas de plantas não transformadas. Por exemplo, as sementes obtidas na maneira descrita acima pode ser plantada e, após um período de desenvolvimento inicial, submetido a uma seleção adequada por pulverização. Uma possibilidade adicional consiste no desenvolvimento das sementes, se apropriado após a esterilização, nas placas ágar usando um agente de seleção adequado de modo que apenas as sementes transformadas pode desenvolver

em plantas. Alternativamente, as plantas transformadas são avaliadas pela presença de um marcador selecionável tal como um descritos acima.

Seguindo a transferência e regeneração de DNA, putativamente as plantas trasnformadas também podem ser avaliadas, por exemplo usando a análise Southern, pela presença do gene de interesse, número de cópia e/ou organização genômica. Alternativamente ou adicionalmente os níveis de expressão de novos DNA introduzidos podem ser monitorados usando a análise Northern e/ou Western, ambas as técnicas sendo bem conhecida às pessoas tendo habilidades comum na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de meios, tal como pelas técnicas de geração clássica ou propagação clonal. Por exemplo, uma primeira geração de planta transformada (ou T1) pode ser auto produzida e transformantes da segunda geração homozigotos (ou T2) selecionados, e as plantas T2 podem ser então propagadas através de técnicas de geração clássica. Os organismos transformados gerados podem obter uma variedade de formas. Por exemplo, estas podem ser quimeras das células transformadas e células não transformadas; transformantes clonais (por exemplo, todas as células transformadas por conter o cassete de expressão); enxertos de tecidos transformados e não transformados (por exemplo, em plantas, um rizoma transformado enxertado a uma muda não transformada).

A presente invenção claramente estende-se a qualquer célula de planta ou planta produzida por qualquer um dos métodos descritos neste, e por todas as partes das plantas e propagules destes. A presente invenção ainda estende-se para abranger a progênie de uma célula transfectada ou transformada primária, tecido, órgão ou total da planta que foi produzida por qualquer um dos métodos anteriormente mencionados, apenas o requerimento sendo que a progênie exibe as mesmas características genotípicas e/ou fenotípicas como aquelas produzidas pela origem dos métodos de acordo

como a invenção.

A invenção também incluem células hospedeiras contendo um ácido nucleico isolado que codifica um polipeptídeo SPL11 como definido anteriormente acima. As células hospedeiras preferidas de acordo como a invenção são células de planta. As plantas hospedeiras pelos ácidos nucleicos ou o vetor usado no método de acordo como a invenção, o cassete de expressão ou construção ou vetor são, à princípio, vantajosamente em todas as plantas, que são capazes de sintetizar os polipeptídeos usados no método da invenção.

Os métodos da invenção são vantajosamente aplicáveis a qualquer planta. As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertence a superfamília Viridiplantae, em particular plantas monocotiledônias e dicotiledônias incluindo forragem ou legumes de forragem, plantas ornamentais, lavouras de alimento, árvores ou arbustos. De acordo com uma forma de realização preferida da presente invenção, a planta é uma planta de cultivo. Exemplos de plantas de cultivo incluem soja, girassol, canola, alfafa, colza, algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta mocotiledônia. Exemplos de planta mocotiledônia incluem cana de açúcar. Mais preferivelmente a planta é um cereal. Exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

A invenção também estende-se à partes colhíveis de uma planta tal como, mas não limitada a sementes, folhas, frutos, flores, talos, raízes, rizomas, tubérculos e bulbos. A invenção além disso relata os produtos derivados, preferivelmente diretamente derivados, a partir de partes colhíveis de uma tal planta, tal como grânulos ou pós secos, óleo, gordura ou ácidos graxos, amido ou proteínas.

De acordo com uma característica preferida da invenção, a expressão modulada é a expressão aumentada. Os métodos para a expressão

aumentada de ácidos nucleicos ou genes, ou produtos dos genes, são bem documentados na técnica e exemplos são fornecidos na seção de definições.

Como mencionado acima, um método preferido para a expressão de modulação (preferivelmente, aumentada) de um ácido nucleico que codifica um polipeptídeo SPL11 é pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo SPL11; entretanto os efeitos da realização do método, isto é intensificação de traços relacionados à produção também podem ser atingidos usando outras técnicas bem conhecidas, incluindo mas não limitada a etiquetagem de ativação de T-DNA, TILLING, recombinação homóloga. Uma descrição desta técnica é fornecida na seção de definições.

A presente invenção também abrange o uso de ácidos nucleicos que codificam os polipeptídeos SPL11 como descrito neste e o uso destes polipeptídeos SPL11 na intensificação de qualquer um dos traços relacionados na produção anteriormente mencionadas em plantas.

Os ácidos nucleicos que codificam os polipeptídeos SPL11 descritos neste, ou os polipeptídeos SPL11 por si mesmos, podem observar o uso nos programas de geração em que um marcador de DNA é identificado, que pode ser geneticamente ligado a um polipeptídeo SPL11 que codifica o gene. Os ácidos nucleicos/genes, ou os polipeptídeos SPL11 por si mesmos podem ser usados para definir um marcador molecular. Este DNA ou marcador de proteína pode ser então usado nos programas de geração para selecionar as plantas tendo os traços relacionados a produção itensificada como definido anteriormente acima nos métodos da invenção.

As variantes alélicas de um ácido nucleico que codifica o polipeptídeo SPL11/gene também pode observar o uso em programas de geração assitidos pelos marcadores. Tais programas de geração algumas vezes requerem a introdução de variação alélica pelo tratamento mutagênico das plantas, usando por exemplo mutagênese EMS; alternativamente, o programa

pode iniciar com uma coleção de variantes alélicas denominadas origem "natural" causadas não intencionalmente. A identificação das variantes alélicas então acontece, por exemplo, por PCR. Este é seguido pela etapa para a seleção de variantes alélicas superiores da sequência em questão e que dão a produção aumentada. A seleção é tipicamente realizada pelo monitoramento do desempenho do desenvolvimento de plantas contendo diferentes variantes alélicas da sequência em questão. O desempenho do desenvolvimento pode ser monitorado em uma estufa ou no campo. Ainda as estapas opcionais incluem o cruzamento de plantas em que a variante alélica superior foi identificada com uma outra planta. Esta pode ser usada, por exemplo, para fabricar uma combinação de características fenotípicas de interesse.

Os ácidos nucleicos que codificam os polipeptídeos SPL11 também podem ser usados como sondas para o mapeamento genético e físico dos genes que estes são uma parte de, e como marcadores para os traços ligados aqueles genes. Tal informação pode ser útil na geração de plantas a fim de desenvolver as linhas com fenotipos desejados. Tal uso de polipeptídeo SPL11 que codifica os ácidos nucleicos requer apenas uma sequência de ácido nucleico em pelo menos 15 nucleotídeos em comprimento. O polipeptídeo SPL11 que codifica os ácidos nucleicos podem ser usados como restrição dos marcadores de poliformismo do comprimento do fragmento (RFLP). Southern blots (Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual) do DNA genômico de planta digerindo a restrição e podem ser sondados com os ácidos nucleicos que codificam o SPL11. Os padrões de ligação resultantes podem ser então submetidos à análise genética usando os programas de computador tal como MapMaker (Lander et.al. (1987) Genomics 1: 174-181) a fim de construir um mapa genético. Além disso, os ácidos nucleicos podem ser usado para sondar Southern blots contendo a restrição do DNA genômico tratado por endonuclease de uma série de indivíduos que representam a origem e

progênie de uma cruzada genética definida. A segregação do polimorfismo do DNA é notado e usado para calcular a posição do ácido nucleico que codifica o polipeptídeo SPL11 no mapa genético previamente obtido usando esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

A produção e uso de sondas derivadas de genes de plantas para o uso no mapeamento genético é descrito em Bernatzky and Tanksley (1986) Plant Mol. Biol. Reporter 4: 37-41. As publicações numerosas descrevem o mapeamento genético de clones de cDNA específicos usando a metodologia resumida acima ou variações destes. Por exemplo, as populações intercruzadas F2, populações cruzadas novamente, aleatoriamente popualções auxiliares, linhas isogênicas próximas, e outras séries de indivíduos podem ser usados para o mapeamento. Tais metodologias são bem conhecidas àqueles habilitados na técnica.

As sondas de ácido nucleico também podem ser usadas para o mapeamento físico (isto é, disposição das sequências em mapas físicos; ver Hoheisel et al. In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346, e refer~encias citadas neste).

Em uma outra forma de realização, as sondas de ácido nucleico podem ser usadas em fluorescência direta no mapeamento de hibridização in situ (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora os métodos correntes do mapeamento FISH favorecem o uso de amplos clones (diversos kb à diversos centos de kb; ver Laan et al. (1995) Genome Res. 5:13-20), melhoramento na sensibilidade pode permitir o desempenho do mapeamento FISH usando sondas mais curtas.

Uma variedade de métodos com base na amplificação do ácido nucleico para o mapeamento físico e genético pode ser realizado usando os ácidos nucleicos. Exemplos incluem amplificação específica por alelos (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332),

ligação específica por alelo (Landegren et al. (1988) Science 241:1077-1080), reações de extensão de nucleotídeo (Sokolov (1990) Nucleic Acids Res. 18:3671), mapeamento híbrida de radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e mapeamento de Happy (Dear and Cook (1989) Nucleic Acids Res. 17:6795-6807). por estes métodos, a sequência de um ácido nucleico é usado para projetar e produzir pares imiciadores para o uso na reação de amplificação ou nas reações de extensão do iniciador. o projeto de tais iniciadores é bem conhecido àqueles habilitados na técnica. Em métodos que utilizam o mapeamento genético com base em PCR, pode ser necessário para identificar as diferenças da sequência de DNA entre os padrões do mapeamento cruzado na região correspondente ao exemplo da sequência de ácido nucleico. Este, entretanto, não é no geral necessário para os métodos do mapeamento.

Os métodos de acordo com uma presente invenção resultam nas plantas tendo os traços relacionados a produção itensificada, como descrito anteriormente mencionadas. Estes traços também podem ser combinados com outros traços vantajosamente econômico, tal como ainda traços que intensificam a produção adicional, tolerância a outras tensões bióticas e abióticas, traços que modificam várias características arquiteturais e/ou características fisiológicas e/ou bioquímicas.

A presente invenção agora será descrita em referência aos aseguintes itens:

1. Um método para a intensificação dos traços relacionados à produção em plantas relativos às plantas de controle, que compreende uma expressão de modulação em uma planta de um ácido nucleico que codifica um polipeptídeo LBD, em que o dito polipeptídeo LBD compreende um domínio DUF206.

2. O método de acordo com o item 1, em que o dito polipeptídeo LBD compreende um ou mais dos seguintes motivos:

(i) Motivo 1: MSCNGCRXLRKGCX (SEQ ID N°: 5),

(ii) Motivo 2: QXXATXFXAKFXGR (SEQ ID N°: 6),

(iii) Motivo 3: FXSLLXEAXG (SEQ ID N°: 7)

3. O método de acordo com o item 1 ou 2, em que a dita expressão modulada é efetuada pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo LBD.

4. O método de acordo com qualquer item antecedente, em que o dito ácido nucleico que codifica um polipeptídeo LBD que codifica qualquer uma das proteínas listadas na Tabela A1 ou é uma porção de um tal ácido nucleico, ou um ácido nucleico capaz de hibridizar com um tal ácido nucleico.

5. O método de acordo com qualquer item antecedente, em que a dita sequência de ácido nucleico codifica um ortólogo ou parálogo de qualquer uma das proteínas dados na Tabela A1.

6. O método de acordo com qualquer item antecedente, em que os ditos traços relacionados a produção itensificada compreende a produção aumentada, preferivelmente a biomassa aumentada e/ou produção aumentada de semente relativas às plantas de controle.

7. O método de acordo com qualquer um dos itens 1 a 6, em que os ditos traços relacionados a produção itensificada são obtidos sob as condições de não tensão.

8. O método de acordo com qualquer um dos itens 1 a 6, em que os ditos traços relacionados a produção itensificada são obtidos sob condições de deficiência de nitrogênio.

9. O método de acordo com qualquer um dos itens 3 a 8, em que o dito ácido nucleico é operavelmente ligado a um promotor constitutivo, preferivelmente a um promotor GOS2, mais preferivelmente a um promotor GOS2 de arroz.

10. O método de acordo com qualquer item antecedente, em

que o dito ácido nucleico que codifica um polipeptídeo LBD é de origem vegetal, preferivelmente de uma planta dicotiledônia, ainda preferivelmente da família Brassicaceae, mais preferivelmente do gênero Arabidopsis, mais preferivelmente de *Arabidopsis thaliana*.

11. Uma molécula de ácido nucleico isolada que compreende qualquer uma das seguintes características:

(i) um ácido nucleico representado pela SEQ ID N°: 69;

(ii) um ácido nucleico ou fragmento do mesmo que é complementar a qualquer uma das SEQ ID N°s dados no (i);

(iii) um ácido nucleico que codifica um polipeptídeo LBD tendo uma ordem aumentada de preferência em pelo menos 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à SEQ ID N°: 70;

(iv) um ácido nucleico capaz da hibridização sob condições estringentes a qualquer um dos ácidos nucleicos dados no (i), (ii) ou (iii) acima.

12. Um polipeptídeo isolado que compreende:

(i) uma sequência de aminoácido tendo uma ordem aumentada de preferência em pelo menos 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou 100 % da identidade de sequência à sequência de aminoácido dada na SEQ ID N°: 70.

(ii) derivados de qualquer uma das sequências de aminoácidos dados no (i).

13. Plantas ou partes destas, incluindo sementes, obtidas pelo método de acordo com qualquer item antecedente, em que as ditas plantas ou partes destas compreendem um ácido nucleico recombinante que codifica um polipeptídeo LBD.

14. Construção que compreende:

(i) ácido nucleico que codifica um polipeptídeo LBD como

definido nos itens 1, 2 ou 12;

(ii) uma ou mais sequências de controle capazes da expressão de condução da sequência do ácido nucleico de (a); e opcionalmente

(iii) uma sequência de terminal de transcrição.

15. Construção de acordo com o item 14, em que uma dita sequência de controle é um promotor constitutivo, preferivelmente um promotor GOS2, mais preferivelmente um promotor GOS2 de arroz.

16. Uso de uma construção de acordo com o item 14 ou 15 em um método para a fabricação das plantas tendo a produção aumentada, particularmente biomassa aumentada e/ou produção aumentada de semente relativas às plantas de controle.

17. Planta, parte da planta ou célula da planta transformada como uma construção de acordo com o item 14 ou 15.

18. Método para a produção de uma planta transgênica tendo a produção aumentada, particularmente a biomassa aumentada e/ou produção aumentada de semente relativas às plantas de controle, que compreende:

(i) a introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo LBD como definido no item 1, 2 ou 12; e

(ii) cultivo da célula de planta sob condições que promovem o desenvolvimento e crescimento vegetal.

19. Planta transgênica tendo a produção aumentada, particularmente a biomassa aumentada e/ou produção aumentada de semente, relativas às plantas de controle, resultando da expressão aumentada de um ácido nucleico que codifica um polipeptídeo LBD como definido no item 1, 2 ou 12, ou uma célula da planta transgênica derivada da dita planta transgênica.

20. Planta transgênica de acordo com o item 13, 17 ou 19, ou uma célula da planta transgênica derivada deste, em que a dita planta é uma

planta de cultivo ou uma monocotiledônia ou um cereal, tal como arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

21. Partes colhíveis de uma planta de acordo com o item 19, em que as ditas partes colhíveis são preferivelmente biomassa de broto e/ou sementes.

22. Produtos derivados de uma planta de acordo com o item 19 e/ou de partes colhíveis de uma planta de acordo com o item 20.

23. Uso de um ácido nucleico que codifica um polipeptídeo LBD na produção aumentada, particularmente na produção aumentada de semente e/ou biomassa de broto em plantas, relativas às plantas de controle.

24. Um método para a intensificação dos traços relacionados de produção em plantas relativos às plantas de controle, que compreende uma expressão de modulação em uma planta de um ácido nucleico que codifica um polipeptídeo JMJC, em que o dito polipeptídeo JMJC compreende um domínio JmjC.

25. O método de acordo com o item 24 em que o dito domínio JmjC é representado pela sequência tendo uma ordem aumentada de preferência em pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, ou mais identidades de sequências a:

(i) SEQ ID N°: 78; e/ou

(ii) um dos domínios JmjC compreendidos nos polipeptídeos JMJC representado pela SEQ ID N°: 84; SEQ ID N°: 86; SEQ ID N°: 96; SEQ ID N°: 98; SEQ ID N°: 104; SEQ ID N°: 108; SEQ ID N°: 110; SEQ ID N°: 112; SEQ ID N°: 114; SEQ ID N°: 116; SEQ ID N°: 118; SEQ ID N°: 120; SEQ ID N°: 122; SEQ ID N°: 124; SEQ ID N°: 128; SEQ ID N°: 130; SEQ ID N°: 132; e SEQ ID N°: 134, aqueles aminoácidos coordenados são dados na Tabela B4

26. O método de acordo com o item 24 e item 25, em que o dito polipeptídeo JMJC que compreende um motivo tendo uma ordem

aumentada de preferência em pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % ou mais identidades de sequências de qualquer uma das:

(i) SEQ ID N°: 79,

(ii) SEQ ID N°: 80,

(iii) SEQ ID N°: 81,

(iv) SEQ ID N°: 82;

27. O método de acordo com o item 24 ou 26, em que a dita expressão modulada é efetuada pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo JMJC.

28. O método de acordo com qualquer um dos itens 24 a 27, em que a dito expressão de modulação é um aumento na expressão.

29. O método de acordo com qualquer um dos itens 24 a 28, em que o dito ácido nucleico que codifica um polipeptídeo JMJC que codifica qualquer uma das proteínas listadas na Tabela B1 ou é uma porção de um tal ácido nucleico, ou um ácido nucleico capaz de hibridizar com um tal ácido nucleico.

30. O método de acordo com um dos itens 24 a 29, em que a dita sequência de ácido nucleico codifica um ortólogo ou parálogo de qualquer uma das proteínas dadas na Tabela B1.

31. O método de acordo com o item 30, em que o dito ácido nucleico codifica a SEQ ID N°: 74.

32. O método de acordo com qualquer um dos itens 24 a 31, em que os ditos traços relacionados a produção itensificada compreende a produção aumentada, preferivelmente índice de colheita aumentada e/ou produção de semente relativas às plantas de controle.

33. O método de acordo com qualquer um dos itens 24 a 32, em que os ditos traços relacionados a produção itensificada compreende o vigor da planta precocemente relativas às plantas de controle.

34. O método de acordo com qualquer um dos itens 24 a 33, em que os ditos traços relacionados a produção itensificada são obtidos sob as condições de não tensão.

35. O método de acordo com qualquer um dos itens 24 a 33, em que os ditos traços relacionados a produção itensificada são obtidos sob brandas condições de desenvolvimento da tensão à aridez.

36. O método de acordo com qualquer um dos itens 24 a 33, em que os ditos traços relacionados a produção itensificada são obtidos sob condições de desenvolvimento da deficiência de nitrogênio.

37. O método de acordo com qualquer um dos itens 27 a 36, em que o dito ácido nucleico é operavelmente ligado a um promotor constitutivo, preferivelmente a um promotor GOS2, mais preferivelmente a um promotor GOS2 de arroz.

38. O método de acordo com qualquer um dos itens 24 a 37, em que o dito ácido nucleico que codifica um polipeptídeo JMJC é de origem vegetal, preferivelmente de uma planta dicotiledônia, ainda preferivelmente da família Brassicaceae, mais preferivelmente do gênero *Arabidopsis,* mais preferivelmente de *Arabidopsis thaliana.*

39. Plantas ou partes destas, incluindo sementes, obtidas pelo método de acordo com qualquer um dos itens 24 a 38, em que as ditas plantas ou partes destas compreendem um ácido nucleico recombinante que codifica um polipeptídeo JMJC.

40. Construção que compreende:

(i) ácido nucleico que codifica um polipeptídeo JMJC como definido nos itens 24 a 26;

(ii) uma ou mais sequências de controle capazes da expressão de condução da sequência do ácido nucleico de (i); e opcionalmente

(iii) uma sequência de terminal de transcrição.

41. Construção de acordo com o item 40, em que uma dita

sequência de controle é um promotor constituitivo, preferivelmente um promotor GOS2, mais preferivelmente um promotor GOS2 de arroz.

42. Uso de uma construção de acordo com o item 40 ou 41 em um método para a fabricação das plantas tendo a produção aumentada dos traços relacionados, particularmente o vigor da planta precocemente e/ou produção aumentada de semente relativas às plantas de controle.

43. Planta, parte da planta ou célula da planta transformada como uma construção de acordo com o item 40 ou 41.

44. Método para a produção de uma planta transgênica tendo a produção aumentada dos traços relacionados, particularmente o vigor da planta precocemente e/ou produção aumentada de semente relativas às plantas de controle, que compreende:

(i) a introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo JMJC como definido no item 24 a 26; e

(ii) cultivo da célula de planta sob condições que promovem o desenvolvimento e crescimento vegetal.

45. Planta transgênica tendo a produção aumentada, particularmente o vigor das mudas de plantas aumentadas e/ou produção aumentada de semente, relativas às plantas de controle, resultando da expressão aumentada de um ácido nucleico que codifica um polipeptídeo JMJC como definido no item 24 a 26, ou uma célula da planta transgênica derivada da dita planta transgênica.

46. Planta transgênica de acordo com o item 39, 43 ou 45, ou uma célula da planta transgênica derivada deste, em que a dita planta é uma planta de cultivo ou uma monocotiledônia ou um cereal, tal como arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

47. Partes colhíveis de uma planta de acordo com o item 46, em que as ditas partes colhíveis são preferivelmente a biomassa de broto e/ou

sementes.

48. Produtos derivados de uma planta de acordo com o item 46 e/ou de partes colhíveis de uma planta de acordo com o item 47.

49. Uso de um ácido nucleico que codifica um polipeptídeo JMJC na produção aumentada, particularmente na produção aumentada de semente e/ou biomassa de broto em plantas, relativas às plantas de controle.

50. Uma molécula de ácido nucleico isolada que compreende qualquer uma das seguintes características:

(i) um ácido nucleico representado pela SEQ ID N°: 169;

(ii) um ácido nucleico ou fragmento do mesmo que é complementar a SEQ ID N°: 169;

(iii) um ácido nucleico que codifica um polipeptídeo JMJC tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à sequência de aminoácido dada na SEQ ID N°: 170;

(iv) um ácido nucleico capaz da hibridização sob as condições estringentes a qualquer um dos ácidos nucleicos dados no (i), (ii) ou (iii) acima.

51. Uma molécula de polipeptídeo isolado que compreende:

(i) uma sequência de aminoácido tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou 100 % ou mais identidades de sequências à sequência de aminoácido dada na SEQ ID N°: 170;

(ii) derivados de qualquer uma das sequências de aminoácidos dados no (i).

52. Um método para a intensificação dos traços relacionados de produção em plantas relativos às plantas de controle, que compreendem uma expressão de modulação em uma planta de um ácido nucleico que codifica uma caseína quinase I (CKI) em que o dito CKI é selecionado da

SEQ ID N°: 174 ou um ortólogo ou parálogo destes.

53. O método de acordo com o item 52, em que a dita expressão modulada é efetuada pela introdução e expressão em uma planta de um ácido nucleico que codifica o o dito polipeptídeo CKI.

54. O método de acordo com o item 52 ou 53, em que o dito ácido nucleico que codifica o dito polipeptídeo CKI é uma porção da SEQ ID N°: 173, ou um ácido nucleico capaz de hibridizar com um tal ácido nucleico.

55. O método de acordo com qualquer um dos itens 52 a 54, em que a dita sequência de ácido nucleico codifica a SEQ ID N°: 174.

56. O método de acordo com qualquer um dos itens 52 a 55, em que os ditos traços relacionados a produção itensificada compreende o vigor precoce aumentado e/ou produção aumentada, preferivelmente a biomassa aumentada e/ou produção aumentada de semente relativas às plantas de controle.

57. O método de acordo com qualquer um dos itens 52 a 56, em que os ditos traços relacionados a produção itensificada são obtidos sob as condições de não tensão.

58. O método de acordo com qualquer um dos itens 52 a 56, em que os ditos traços relacionados a produção itensificada são obtidos sob as condições de tensão abiótica.

59. O método de acordo com o item 58, em que as condições de tensão abiótica são selecionados de um ou mais de: conditions de tensão à aridez, condições de tensão ao sal, e condições da deficiência de nitrogênio.

60. O método de acordo com qualquer um dos itens 53 a 59, em que o dito ácido nucleico é operavelmente ligado a um promotor específico por semente.

61. O método de acordo com qualquer um dos itens 53 a 60, em que o dito promotor específico por semente é um promotor WSI18, preferivelmente a um promotor WSI18 de arroz.

62. O método de acordo com qualquer um dos itens 52 a 61, em que o dito ácido nucleico que codifica um polipeptídeo CKI é de origem vegetal.

63. O método de acordo com qualquer um dos itens 52 a 62, em que a dita origem da planta é de preferivelmente uma planta dicotiledônia, ainda preferivelmente da família Solanaceae, mais preferivelmente do gênero *Nicotiana tabacum.*

64. Plantas ou partes destas, incluindo sementes, obtidas pelo método de acordo com qualquer um dos itens 52 a 63, em que as ditas plantas ou partes destas compreendem um ácido nucleico recombinante que codifica um polipeptídeo CKI.

65. Construção que compreende:

(i) ácido nucleico que codifica um polipeptídeo CKI como definido no item 52;

(ii) uma ou mais sequências de controle capazes da expressão de condução da sequência do ácido nucleico de (a); e opcionalmente

(iii) uma sequência de terminal de transcrição.

66. Construção de acordo com o item 65, em que uma dita sequência de controle é um promotor específico de semente, preferivelmente um promotor WSI18.

67. Construção de acordo com o item 65, em que um dito promotor específico de semente é um promotor WSI18, mais preferivelmente um promotor WSI18 de arroz.

68. Uso de uma construção de acordo com qualquer um dos itens 65 a 67 em um método para a fabricação das plantas tendo a produção dos traços relacionadas aumentadas, particularmente o vigor precoce aumentado, biomassa aumentada e/ou produção aumentada de semente relativas às plantas de controle.

69. Planta, parte da planta ou célula da planta transformada

como uma construção de acordo com qualquer um dos itens 65 a 67.

70. Método para a produção de uma planta transgênica tendo a produção aumentada, particularmente a biomassa aumentada e/ou produção aumentada de semente relativas às plantas de controle, que compreende:

(i) a introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo CKI como definido no item 52; e

(ii) cultivo da célula de planta sob as condições que promovem o desenvolvimento e crescimento vegetal.

71. Planta transgênica tendo a produção dos traços relacionados aumentados, particularmente o vigor precoce aumentado, a biomassa aumentada e/ou produção aumentada de semente, relativas às plantas de controle, resultando da expressão aumentada de um ácido nucleico que codifica um polipeptídeo CKI como definido no item 52, ou uma célula da planta transgênica derivada da dita planta transgênica.

72. Planta transgênica de acordo com o item 64, 69 ou 71, ou uma célula da planta transgênica derivada deste, em que a dita planta é uma planta de cultivo ou uma monocotiledônia ou um cereal, tal como arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

73. Partes colhíveis de uma planta de acordo com o item 72, em que as ditas partes colhíveis são preferivelmente a biomassa de broto e/ou sementes.

74. Produtos derivados de uma planta de acordo com o item 72 e/ou de partes colhíveis de uma planta de acordo com o item 73.

75. Uso de um ácido nucleico que codifica um polipeptídeo CKI na produção dos traços relacionadas aumentados, particularmente no aumento de um ou mais do vigor precoce, produção de semente e biomassa de broto em plantas, relativos às plantas de controle.

76. Um método para a intensificação dos traços relacionados de produção, preferivelmente intensificando os traços relacionados à

produção de semente em plantas relativas às plantas de controle, que compreende a modulação, preferivelmente a expressão aumentada em uma planta de uma sequência de ácido nucleico que codifica um polipeptídeo de prolongamento-homeodomínio de homeodomínio de planta (PHDf-HD), que o polipeptídeo PHDf-HD compreende: (i) um domínio tendo em pelo menos 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais da identidade de aminoácidos de sequência a um domínio de prolongamento-homeodomínio de planta / zíper de leucina como representado pela SEQ ID N°: 233; e (ii) um domínio tendo em pelo menos 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais da identidade de aminoácidos de sequência a um homeodomínio (HD) como representado pela SEQ ID N°: 234, e opcionalmente a seleção pelas plantas tendo os traços relacionados a produção itensificada.

77. O método de acordo com o item 76, em que o dito polipeptídeo PHDf-HD compreende: (i) um domínio PHD como representado por PFAM00628; e (ii) um HD como representado por PFAM00046.

78. O método de acordo com o item 76 ou 77, em que o dito polipeptídeo PHDf-HD, quando usado na construção de uma árvore filogenética HD, tal como um descrito na Figura 13, os grupos com o grupo PHDf-HD dos polipeptídeos que compreendem a sequência de polipeptídeo como representado pela SEQ ID N°: 180, antes do que com qualquer outro grupo HD.

79. O método de acordo com qualquer um dos itens 76 a 78, em que o dito polipeptídeo PHDf-HD tem uma ordem aumentada de preferência em pelo menos 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais identidades de sequências ao polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180 ou qualquer um da sequência de polipeptídeos dados na Tabela D1 neste.

80. O método de acordo com qualquer um dos itens 76 a 79, em que a dita sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD é representado por qualquer uma da sequência do ácido nucleico SEQ ID N° dado na Tabela D1 ou uma porção deste, ou uma sequência capaz de hibridizar com qualquer uma da sequência dos ácidos nucleicos SEQ ID N° dados na Tabela D1.

81. O método de acordo com qualquer um dos itens 76 a 80, em que a dita sequência de ácido nucleico codifica um ortólogo ou parálogo de qualquer uma da SEQ ID N° dada na Tabela D1.

82. O método de acordo com qualquer um dos itens 76 a 81, em que o dito modulado, preferivelmente a expressão aumentada é efetuada por qualquer um ou mais de: Etiquetagem de ativação de T-DNA, TILLING, ou recombinação homóloga.

83. O método de acordo com qualquer um dos itens 76 a 82, em que a dita expressão aumentada é efetuada pela introdução e expressão em uma planta de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD.

84. O método de acordo com qualquer um dos itens 76 a 83, em que os ditos traços relacionados à produção são os traços relacionados à semente de produção, que compreendem um ou mais de: (i) número aumentado de panícula primária; (ii) peso da semente total aumentado por planta; (iii) número aumentado de sementes (enchidas); (iv) TKW aumentado; ou (v) índice de colheita aumentada.

85. O método de acordo com qualquer um dos itens 76 a 84, em que a dita sequência de ácido nucleico é operavelmente ligado a um promotor constitutivo, preferivelmente a um promotor constitutivo de planta, mais preferivelmente a um promotor GOS2, mais preferivelmente a um promotor GOS2 de arroz.

86. O método de acordo com qualquer um dos itens 76 a 85,

em que a dita sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD é de origem vegetal, preferivelmente a partir de uma planta mocotiledônia, ainda preferivelmente da família Poacae mais preferivelmente do gênero Oryza, mais preferivelmente de *Oryza sativa.*

87. Plantas, partes destas (incluindo sementes), ou células de planta obtidas pelo método de acordo com qualquer um dos itens 76 a 86, em que a dita planta, parte ou célula destas compreendem um transgene do ácido nucleico isolado que codifica um polipeptídeo PHDf-HD operavelmente ligado a um promotor constitutivo de planta.

88. Construção que compreende:

(i) a sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD como definido em qualquer um dos itens 76 a 81;

(ii) uma ou mais sequências de controle capazes da expressão de condução da sequência do ácido nucleico de (i); e opcionalmente

(iii) uma sequência de terminal de transcrição.

em que em pelo menos uma da sequência de controle é um promotor constitutivo de planta, preferivelmente um promotor GOS2.

89. Uso de uma construção de acordo com o item 87 em um método para a fabricação das plantas tendo os traços relacionados a produção itensificada relativos às plantas de controle, no qual os traços relacionados a produção itensificada, preferivelmente os traços relacionados à semente da produção itensificada, são um ou mais de: (i) número aumentado de panícula primária; (ii) peso da semente total aumentado por planta; (iii) número aumentado de sementes (enchidas); (iv) TKW aumentado; ou (v) índice de colheita aumentada.

90. Planta, parte da planta ou célula da planta transformada como uma construção de acordo com o item 87 ou 88.

91. O método para a produção de plantas transgênicas tendo os traços relacionados a produção itensificada relativos às plantas de controle,

que compreende:

(i) a introdução e expressão em uma planta, parte da planta ou célula de planta, uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD como definido em qualquer um dos itens 76 a 81, sob o controle do promotor constitutivo de planta; e

(ii) cultivo da célula de planta sob as condições que promovem o desenvolvimento e crescimento vegetal.

92. Planta transgênica tendo os traços relacionados a produção itensificada, preferivelmente os traços relacionados à semente da produção itensificada, relativos às plantas de controle, resultando da modulação, preferivelmente a expressão aumentada de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD como definido em qualquer um dos itens 76 a 81, operavelmente ligado a um promotor constitutivo de planta, ou uma célula da planta transgênica derivada da dita planta transgênica.

93. Planta transgênica de acordo com o item 87, 90 ou 92, em que a dita planta é uma planta de cultivo ou uma monocotiledônia ou um cereal, tal como arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias, ou uma célula da planta transgênica derivada da dita planta transgênica.

94. Partes colhíveis que compreendem uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD de uma planta de acordo com o item 93, em que as ditas partes colhíveis são preferivelmente sementes.

95. Produtos derivados de uma planta de acordo com o item 93 e/ou de partes colhíveis de uma planta de acordo com o item 94.

96. Uso de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD como definido em qualquer um dos itens 76 a 81 em intensificação de traços relacionados à produção em plantas, preferivelmente na intensificação dos traços relacionados à semente da produção, que compreende um ou mais de: (i) número aumentado de panículas primárias;

(ii) peso da semente total aumentado por planta; (iii) número aumentado de sementes (enchidas); (iv) TKW aumentado; ou (v) índice de colheita aumentada.

97. Uma molécula de ácido nucleico isolada que compreende qualquer uma das seguintes características:

(i) um ácido nucleico representado pela SEQ ID N°: 242;

(ii) um ácido nucleico ou fragmento do mesmo que é complementar a SEQ ID N°: 242;

(iii) um ácido nucleico que codifica um polipeptídeo JMJC tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à sequência de aminoácido dada na SEQ ID N°: 243;

(iv) um ácido nucleico capaz da hibridização sob as condições estringentes a qualquer um dos ácidos nucleicos dados no (i), (ii) ou (iii) acima.

98. Uma molécula de polipeptídeo isolado que compreende:

(i) uma sequência de aminoácido tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou 100 % ou mais identidades de sequências à sequência de aminoácido dada na SEQ ID N°: 243;

(ii) derivados de qualquer uma das sequências de aminoácidos dados no (i).

99. Um método para a intensificação dos traços relacionados de produção em plantas relativas às plantas de controle, que compreendem uma expressão de modulação em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante ao bHLH-11, em que o dito polipeptídeo semelhante ao bHLH-11 compreende um Domínio Hélice-Laço-Hélice.

100. O método de acordo com o item 99, em que o dito polipeptídeo semelhante ao bHLH-11 compreende um ou mais dos seguintes

motivos:

(i) Motivo 1 (SEQ ID N°: 246);

(ii) Motivo 2 (SEQ ID N°: 247);

(iii) Motivo 3 (SEQ ID N°: 248);

(iv) Motivo 4 (SEQ ID N°: 249);

(v) Motivo 5 (SEQ ID N°: 250);

(vi) Motivo 6 (SEQ ID N°: 251);

(vii) Motivo 7 (SEQ ID N°: 252).

101. O método de acordo com o item 99 ou 100, em que a dita expressão modulada é efetuada pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante ao bHLH-11.

102. O método de acordo com qualquer um dos itens 99 a 101, em que o dito ácido nucleico que codifica um polipeptídeo semelhante ao bHLH-11 que codifica qualquer uma das proteínas listadas na Tabela E1 ou é uma porção de um tal ácido nucleico, ou um ácido nucleico capaz de hibridizar com um tal ácido nucleico.

103. O método de acordo com qualquer um dos itens 99 a 102, em que a dita sequência de ácido nucleico codifica um ortólogo ou parálogo de qualquer uma das proteínas dadas na Tabela E1.

104. O método de acordo com qualquer um dos itens 99 a 103, em que os ditos traços relacionados a produção itensificada compreende a produção aumentada, preferivelmente produção aumentada de semente relativos às plantas de controle.

105. O método de acordo com qualquer um dos itens 99 a 104, em que os ditos traços relacionados a produção itensificada são obtidos sob as condições de não tensão.

106. O método de acordo com qualquer um dos itens 101 a 105, em que o dito ácido nucleico é operavelmente ligado a um promotor constitutivo, preferivelmente a um promotor GOS2, mais preferivelmente a

um promotor GOS2 de arroz.

107. O método de acordo com qualquer um dos itens 99 a 106, em que o dito ácido nucleico que codifica um polipeptídeo semelhante ao bHLH-11 é de origem vegetal, preferivelmente a partir de uma planta mocotiledônia, ainda preferivelmente da família Poaceae, mais preferivelmente do gênero *Triticum,* mais preferivelmente de *Triticum aestivum.*

108. Plantas ou partes destas, incluindo sementes, obtidas pelo método de acordo com qualquer um dos itens 99 a 107, em que as ditas plantas ou partes destas compreende um ácido nucleico recombinante que codifica um polipeptídeo semelhante ao bHLH-11.

109. Construção que compreende:

(i) ácido nucleico que codifica um polipeptídeo semelhante ao bHLH-11 como definido nos itens 99 ou 100;

(ii) uma ou mais sequências de controle capazes da expressão de condução da sequência do ácido nucleico de (i); e opcionalmente

(iii) uma sequência de terminal de transcrição.

110. Construção de acordo com o item 109, em que uma dita sequência de controle é um promotor constitutivo, preferivelmente um promotor GOS2, mais preferivelmente um promotor GOS2 de arroz.

111. Uso de uma construção de acordo com o item 109 ou 110 em um método para a fabricação das plantas tendo a produção aumentada, particularmente a produção aumentada de semente relativas às plantas de controle.

112. Planta, parte da planta ou célula da planta transformada como uma construção de acordo com o item 109 ou 110.

113. Método para a produção de uma planta transgênica tendo a produção aumentada, particularmente a produção aumentada de semente relativas às plantas de controle, que compreende:

(i) a introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante ao bHLH-11 como definido no item 99 ou 100; e

(ii) cultivo da célula de planta sob as condições que promovem o desenvolvimento e crescimento vegetal.

114. Planta transgênica tendo a produção aumentada, particularmente a produção aumentada de semente, relativas às plantas de controle, resultando da expressão modulada de um ácido nucleico que codifica um polipeptídeo semelhante ao bHLH-11 como definido no item 99 ou 100, ou uma célula da planta transgênica derivada da dita planta transgênica.

115. Planta transgênica de acordo com o item 108, 112 ou 114, ou uma célula da planta transgênica derivada deste, em que a dita planta é uma planta de cultivo ou uma monocotiledônia ou um cereal, tal como arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo, espécie de trigo, espelta, *secale, einkorn, teff, milo* e aveias.

116. Partes colhíveis de uma planta de acordo com o item 115, em que as ditas partes colhíveis são preferivelmente sementes.

117. Produtos derivados de uma planta de acordo com o item 115 e/ou de partes colhíveis de uma planta de acordo com o item 116.

118. Uso de um ácido nucleico que codifica um polipeptídeo semelhante ao bHLH-11 na produção aumentada, particularmente na produção aumentada de semente em plantas, relativos às plantas de controle.

119. Um método para a intensificação dos traços relacionados de produção em plantas relativos às plantas de controle, que compreende uma expressão de modulação em uma planta de um ácido nucleico que codifica um ASR, em que o dito ASR é representado pela SEQ ID N°: 397 ou um ortólogo ou parálogo destes.

120. O método de acordo com o item 119, em que a dita

expressão modulada é efetuada pela introdução e expressão em uma planta de um ácido nucleico que codifica o dito polipeptídeo ASR.

121. O método de acordo com o item 119 ou 120, em que o dito ácido nucleico que codifica o dito polipeptídeo ASR é uma porção da SEQ ID N°: 396, ou um ácido nucleico capaz de hibridizar com um tal ácido nucleico.

122. O método de acordo com qualquer um dos itens 119 a 121, em que a dita sequência de ácido nucleico codifica a SEQ ID N°: 397, ou um ortólogo ou parálogo destes.

123. O método de acordo com qualquer um dos itens 119 a 122, em que os ditos traços relacionados a produção itensificada compreende a produção aumentada, preferivelmente a produção aumentada de semente relativas às plantas de controle.

124. O método de acordo com qualquer um dos itens 119 a 123, em que os ditos traços relacionados a produção itensificada são obtidas sob as condições de não tensão.

125. O método de acordo com qualquer um dos itens 120 a 124, em que o dito ácido nucleico é operavelmente ligado a um promotor constitutivo, preferivelmente a um promotor GOS2, mais preferivelmente a um promotor GOS2 de arroz.

126. O método de acordo com qualquer um dos itens 119 a 125, em que o dito ácido nucleico que codifica um polipeptídeo ASR é de origem vegetal, preferivelmente de uma planta dicotiledônia, ainda preferivelmente da família Poaceae, mais preferivelmente do gênero Oryza, mais preferivelmente de *Oryza sativa.*

127. Plantas ou partes destas, incluindo sementes, obtidas pelo método de acordo com qualquer um dos itens 119 a 126, em que as ditas plantas ou partes destas compreendem um ácido nucleico recombinante que codifica um polipeptídeo ASR.

128. Construção que compreende:

(i) ácido nucleico que codifica um polipeptídeo ASR como definido no item 119 ou qualquer um da SEQ ID N°: 401, 403, 405, 407, 409, 411, 413, 415 e 417;;

(ii) uma ou mais sequências de controle capazes da expressão de condução da sequência do ácido nucleico de (a); e opcionalmente

(iii) uma sequência de terminal de transcrição.

129. Construção de acordo com o item 128, em que uma dita sequência de controle é um promotor constitutivo, preferivelmente um promotor GOS2, mais preferivelmente um promotor GOS2 de arroz.

130. Uso de uma construção de acordo com qualquer um dos itens 128 ou 129 em um método para a fabricação das plantas tendo a produção dos traços relacionadas aumentadas, particularmente a produção aumentada de semente relativos às plantas de controle.

131. Planta, parte da planta ou célula da planta transformada como uma construção de acordo com qualquer um dos itens 128 ou 129.

132. Método para a produção de uma planta transgênica tendo a produção aumentada, particularmente a biomassa aumentada e/ou produção aumentada de semente relativas às plantas de controle, que compreende:

(i) a introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo ASR como definido no item 119; e

(ii) cultivo da célula de planta sob as condições que promovem o desenvolvimento e crescimento vegetal.

133. Planta transgênica tendo a produção dos traços relacionados aumentados, particularmente a produção aumentada de semente, relativos às plantas de controle, resultando da expressão aumentada de um ácido nucleico que codifica um polipeptídeo ASR como definido no item 119, ou uma célula da planta transgênica derivada da dita planta transgênica.

134. Planta transgênica de acordo com o item 127, 131 ou 133,

ou uma célula da planta transgênica derivada deste, em que a dita planta é uma planta de cultivo ou uma monocotiledônia ou um cereal, tal como arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

135. Partes colhíveis de uma planta de acordo com o item 134, em que as ditas partes colhíveis são preferivelmente sementes.

136. Produtos derivados de uma planta de acordo com o item 134 e/ou de partes colhíveis de uma planta de acordo com o item 135.

137. Uso de um ácido nucleico que codifica um polipeptídeo ASR nos traços relacionados à produção aumentada, particularmente na produção aumentada de semente em plantas, relativos às plantas de controle.

138. Uma molécula de ácido nucleico isolada que compreende qualquer uma das seguintes características:

(i) um ácido nucleico representado pela qualquer um da SEQ ID N°: 401, 403, 405, 407; 409, 411, 413, 415 e 417;

(ii) o complemento de um ácido nucleico representado pela qualquer um da SEQ ID N°: 401, 403, 405, 407, 409, 411, 413, 415 e 417;

(iii) um ácido nucleico que codifica um polipeptídeo ASR tendo uma ordem aumentada de preferência em pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à sequência de aminoácido representada pela qualquer um da SEQ ID N°: 402, 404, 406, 408, 410, 412, 414, 416 e 418.

139. Um polipeptídeo isolado que compreende:

(i) uma sequência de aminoácido representada pela qualquer um da SEQ ID N°: 402, 404, 406, 408, 410, 412, 414, 416 e 418;

(ii) uma sequência de aminoácido tendo uma ordem aumentada de preferência em pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à sequência de aminoácido representado por qualquer uma da SEQ ID N°: 402, 404, 406, 408, 410, 412, 414, 416 e 418.

(iii) derivados de qualquer uma das sequências de aminoácidos dados no (i) ou (ii) acima.

140. Um método para a intensificação dos traços relacionados de produção em plantas relativos às plantas de controle, que compreende uma expressão de modulação em uma planta de um ácido nucleico que codifica um polipeptídeo SPL11, em que o dito polipeptídeo SPL11 compreende um domínio SBP tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97 % ou mais identidades de sequências a qualquer uma das SEQ ID N°s: 456 a SEQ ID N°: 468 e SEQ ID N°: 478.

141. O método de acordo com o item 140 em que o dito polipeptídeo SPL11 em adição ao domínio SBP compreende qualquer um ou mais dos seguintes motivos conservados:

(i) Motivo 1 como representado pela SEQ ID N°: 469 em que qualquer substituição de aminoácido conservativo e/ou 1 ou 2 substituição não conservativa são permitidas;

(ii) Motivo 2 como representado pela SEQ ID N°: 470 em que qualquer mudança é permitida, fornecida que em pelo menos 4 aminoácidos tem uma cadeia secundária polar, preferivelmente serina ou treonina, e fornecido que o domínio é localizado na extremidade do terminal N do domínio SBP;

(iii) Motivo 3 como representado pela SEQ ID: 471 em que 1 ou 2 más combinações são permitidas;

(iv) Motivo 4 como representado pela SEQ ID: 472 em que 1, 2 ou 3 más combinações são permitidas.

142. O método de acordo com o item 140 ou 141, em que a dita expressão modulada é efetuada pela introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo SPL11 como definido no item 140 ou 141.

143. O método de acordo com qualquer um dos itens 140 a

142, em que a dita expressão modulada é a expressão aumentada.

144. O método de acordo com qualquer um dos itens 140 a 143, em que o dito ácido nucleico que codifica um polipeptídeo SPL11 que codifica qualquer uma das proteínas listadas na Tabela G1 ou é uma porção de um tal ácido nucleico, ou um ácido nucleico capaz de hibridizar com um tal ácido nucleico.

145. O método de acordo com qualquer um dos itens 140 a 144, em que a dita sequência de ácido nucleico codifica um ortólogo ou parálogo de qualquer uma das proteínas dados na Tabela G1.

146. O método de acordo com o item 145, em que o dito ácido nucleico codifica a SEQ ID N°: 428.

147. O método de acordo com qualquer um dos itens 140 a 146, em que os ditos traços relacionados a produção itensificada compreende a produção aumentada, preferivelmente o peso da semente total aumentado, número de sementes enchidas, número de sementes ou florzinhas por panícula, peso de núcleo milhar, taxa de enchimento da semente, e/ou índice de colheita relativos às plantas de controle.

148. O método de acordo com qualquer um dos itens 140 a 147, em que os ditos traços relacionados a produção itensificada compreende o vigor vegetal (mudas) precocemente relativos às plantas de controle.

149. O método de acordo com qualquer um dos itens 140 a 147, em que os ditos traços relacionados a produção itensificada são obtidos sob as condições de não tensão.

150. O método de acordo com qualquer um dos itens 140 a 147, em que os ditos traços relacionados a produção itensificada são obtidos sob brandas condições de desenvolvimento da tensão à aridez.

151. O método de acordo com qualquer um dos itens 140 a 147, em que os ditos traços relacionados a produção itensificada são obtidos sob as condições de desenvolvimento da deficiência de nitrogênio.

152. O método de acordo com qualquer um dos itens 140 a 151, em que o dito ácido nucleico é operavelmente ligado a um promotor constitutivo, preferivelmente a um promotor GOS2, mais preferivelmente a um promotor GOS2 de arroz.

153. O método de acordo com qualquer um dos itens 140 a 151, em que o dito ácido nucleico é operavelmente ligado a um promotor específico de semente, preferivelmente a um promotor WSI18, mais preferivelmente a um promotor WSI18 de arroz.

154. O método de acordo com qualquer um dos itens 140 a 153, em que o dito ácido nucleico que codifica um polipeptídeo SPL11 é de origem vegetal, preferivelmente de uma planta dicotiledônia, ainda preferivelmente da família Brassicaceae, mais preferivelmente do gênero *Arabidopsis,* mais preferivelmente de *Arabidopsis thaliana.*

155. Plantas ou partes destas, incluindo sementes, obtidas pelo método de acordo com qualquer um dos itens 140 a 154, em que as ditas plantas ou partes destas compreendem um ácido nucleico recombinante que codifica um polipeptídeo SPL11.

156. Uma molécula de ácido nucleico isolada que compreende:

(i) um ácido nucleico representado pela SEQ ID N°: 448;

(ii) o complemento de um ácido nucleico representado pela SEQ ID N°: 448;

(iii) um ácido nucleico que codifica um polipeptídeo SPL11 tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à sequência de aminoácido representada pela SEQ ID N°: 449 e tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à SEQ ID N°: 465:

SYCQVEGCRTDLSSAKDYHRKHRVCEPHSKAPKVVVA

GLERRFCQQCSRFHG LAEFDQKKKSCRRRLNDHNARRRKPQPEAL;

(iv) um ácido nucleico que hidrodiza sob as condições estringentes à SEQ ID N°: 448.

157. Um polipeptídeo isolado que compreende:

(i) uma sequência de aminoácido representada pela SEQ ID N°: 449;

(ii) uma sequência de aminoácido tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à sequência de aminoácido representada pela SEQ ID N°: 449 e tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidades de sequências à SEQ ID N°: 465:

SYCQVEGCRTDLSSAKDYHRKHRVCEPHSKAPKVVVA GLERRFCQQCSRFHGLAEFDQKKKSCRRRLNDHNARRRKPQPEAL.

(iii) derivados de qualquer uma das sequências de aminoácidos dados no (i) ou (ii) acima.

158. Construção que compreende:

(i) ácido nucleico que codifica um polipeptídeo SPL11;

(ii) uma ou mais sequências de controle capazes da expressão de condução da sequência do ácido nucleico de (i); e opcionalmente

(iii) uma sequência de terminal de transcrição

159. Construção de acordo com o item 158 em que o dito ácido nucleico que codifica um polipeptídeo SPL11 é um ácido nucleico de acordo com o item 156.

160. Construção de acordo com o item 158 ou 159 em que uma dita sequência de controle é um promotor constitutivo, preferivelmente um promotor GOS2, mais preferivelmente um promotor GOS2 de arroz.

161. Construção de acordo com o item 158 ou 159 em que uma dita sequência de controle é um promotor específico de semente,

preferivelmente um promotor WSI18, mais preferivelmente um promotor WSI18 de arroz.

162. Uso de uma construção de acordo com o item 158 to 161 em um método para a fabricação das plantas tendo a produção aumentada dos traços relacionados; particularmente a produção aumentada de semente relativos às plantas de controle e/ou vigor de mudas ou plantas precocemente.

163. Planta, parte da planta ou célula da planta transformada como uma construção de acordo com o item 158 a 161.

164. Método para a produção de uma planta transgênica tendo a produção aumentada dos traços relacionados, particularmente a produção aumentada de semente e/ou vigor de mudas ou plantas precocemente relativos às plantas de controle, que compreende:

(i) a introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo SPL11; e

(ii) cultivo da célula de planta sob as condições que promovem o desenvolvimento e crescimento vegetal.

165. Planta transgênica tendo a produção aumentada, particularmente o vigor das mudas de plantas aumentadas e/ou produção aumentada de semente, relativos às plantas de controle, resultando da expressão aumentada de um ácido nucleico que codifica um polipeptídeo SPL11, ou uma célula da planta transgênica derivada da dita planta transgênica.

166. Planta transgênica de acordo com o item 155, 163 ou 165, ou uma célula da planta transgênica derivada deste, em que a dita planta é uma planta de cultivo dicotiledônia tal como soja, algodão ou canola ou uma planta de cultivo monocotiledônia ou um cereal, tal como arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

167. Partes colhíveis de uma planta de acordo com o item 166, em que as ditas partes colhíveis são preferivelmente biomassa de broto, flores

e/ou sementes.

168. Produtos derivados de uma planta de acordo com o item 166 e/ou de partes colhíveis de uma planta de acordo com o item 167.

169. Uso de um ácido nucleico que codifica um polipeptídeo SPL11 na produção aumentada, particularmente na produção aumentada de semente e/ou biomassa de broto em plantas, relativos às plantas de controle.

Descrição das figuras

A presente invenção será descrita com referência às seguintes figuras em que:

I. Domínio LOB que compreende a proteína (LOB: Fronteiras orgânicas laterais)

Figura 1 representa uma sequência da SEQ ID N°: 2, com o domínio DUF260 mostrado em negrito, os motivos conservados 1, 2 e 3 são sublinhados e os resíduos Cys conservados (Motivo da SEQ ID N°: 9) são mostrados em itálico.

Figura 2 representa um alinhamento múltiplo da sequências de várias proteínas LBD úteis nos métodos da presente invenção.

Figura 3 mostra uma árvore filogenética de classe II e classe I das proteínas LBD. A sequência da SEQ ID N°: 2 é representada pela AtLBD37.

Figura 4 representa o vetor binário para a expressão aumentada na *Oryza sativa* de um ácido nucleico que codifica LBD sob o controle de um promotor GOS2 de arroz (pGOS2)

Figura 5 detalha os exemplos das sequências LBD úteis na realização de métodos de acordo com uma presente invenção.

II. Polipeptídeo JMJC (JUMONJI-C)

Figura 6 representa uma sequência de aminoácido e estrutura de domínio da SEQ ID N°: 74. Motivos conservados e domínios são indicados. O domínio JmjC é indicado em negrito. Motivos conservados da

SEQ ID N°: 79, SEQ ID N°: 80 e SEQ ID N°: 81 são indicados por retângulos de linha simples, duplos e triplos respectivamente. Sublinhados são os resíduos de aminoácidos T (Treonina) e K (Lisina) que tipicamente participam na coordenação de 2-Oxogutarato. A letra H indica o resíduo de aminoácido de Histidina que coordena o íon ferroso.

Figura 7 representa um alinhamento múltiplo de proteínas JMJC. A origem da proteína é indicada pelos dois dígitos alfa numéricos no nome, At: *Arabidopsis thaliana*, Pp: Populus trichocarpa, Os: *Oryza sativa*, Ot: streococcus tauri, Ce: Caenorhabditis elegans, Hs: Homo sapiens. A posição dos resíduos de aminoácidos conservados e domínios correspondentes aqueles descritos na Figura 6 é indicado. Uma sequência de consenso é dada. Os aminoácidos altamente conservados nas sequências de consenso são dados; espaços em branco representam qualquer aminoácido.

Figura 8 mostra uma árvore filogenética de polipeptídeos JMJC. A seta mostra a SEQ ID N°: 74. Outros polipeptídeos JMJC são nomeados usando o Número de Acessão Genbank. O grupo I (GI) compreende as proteínas de origem vegetal; o grupo II (GII) compreende as proteínas de origem não vegetal.

Figura 9 representa o vetor binário para a expressão aumentada em *Oryza sativa* da SEQ ID N°: 73 sob o controle de um promotor GOS2 de arroz (pGOS2).

Figura 10 detalha os exemplos das sequências JMJC úteis na realização de métodos de acordo com uma presente invenção.

III. Caseína de quinase I

Figura 11 representa o vetor binário para a expressão aumentada em *Oryza sativa* de um ácido nucleico que codifica GRP sob o controle de um promotor WSI18 de arroz (pWSI18::GRP)

Figura 12 detalha os exemplos das sequências GRP úteis na realização de métodos de acordo com uma presente invenção.

IV. Polipeptídeo de prolongamento-homeodomínio de homeodomínio de planta (PHDf-HD)

Figura 13 representa uma árvore de ligação próxima construída após um alinhamento de todos os fatores de transcrição pertencentes à família HD (baixados a partir dos bancos de dados TFDB de arroz hospedado no servidor da University of Potsdam) e todos da sequência de polipeptídeos da Tabela D1 (quando o comprimento total), realizado usando o algorítmo Clustal (1,83) do alinhamento progressivo, usando valores default. O grupo de interesse, que compreende os dois parálogos de arroz (SEQ ID N°: 180 ou Os02g05450.1, e a SEQ ID N°: 202 ou Os06g12400.1) foram circulados. Qualquer polipeptídeo diminuído dentro deste grupo HD (após uma nova etapa de alinhamento múltiplo como descrito anteriomente acima) é considerado ser útil na realização de métodos da invenção como descrito neste.

Figura 14 representa um desenho de um polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180, que compreende uma ou mais das seguintes características: um sinal de localização nuclear predito (NLS), um zíper de leucina (ZIP), um dedo PHD (PHDf, Pfam00628), uma extensão ácida, duas extensões básicas, um homeodomínio (HD, Pfam00046).

Figura 15 mostra os logotipos da sequência do homeodomínio (HD) dos polipeptídeos PHDf-HD da Tabela D1, onde a altura total das pilhas indicam uma conservação de sequência em que posição, enquanto a altura dos símbolos dentro da pilha indica a frequência relativa de cada amino ou ácido nucleico em que posição. O HD como representado pela SEQ ID N°: 234, e compreendido na SEQ ID N°: 180, é de acordo com um logotipo da sequência como representado nesta Figura.

Figura 16 mostra a produção dos gráficos do algorítmo COILS que prediz os dois domínios do espiral enrolado na metade do terminal N do polipeptídeo como representado pela SEQ ID N°: 180. O eixo X representa as

coordenadas do resíduo de aminoácido, a probabilidade do eixo Y (variando de 0 a 1) que um domínio de espiral enrolado está presente, e as três linhas, as três janelas (14, 21, 28) examinadas.

Figura 17 mostra um alinhamento da sequência múltipla do CLUSTAL W (1,83) dos polipeptídeos PHDf-HD a partir da Tabela D1 (quando o comprimento total), onde um número de características são identificadas. A partir do terminal N ao terminal C dos polipeptídeos são: (i) um sinal de localização nuclear predito (NLS); (ii) um zíper de leucina (ZIP), com quatro *heptads* (em compartimentos, em que usualmente uma leucina (ocasionalmente uma isoleucina, uma valina, ou uma metionina) aparece a cada sétimo aminoácido); (iii) um dedo PHD (PHDf), com o C4HC3 típico (quatro cisteínas, uma histidina, três cisteínas) com um espaçamento de cisteína característico; (iv) uma extensão ácida (rico em aminoácidos ácidos D e E); (v) extensões básicas (rica em aminoácidos básicos K e R); (vi) um homeodomínio (HD).

Figura 18 mostra o vetor binário para o modulado, preferivelmente a expressão aumentada em *Oryza sativa* de uma sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD sob o controle de um promotor GOS2 de arroz (pGOS2)

Figura 19 detalha os exemplos das sequências PHDf-HD úteis na realização de métodos de acordo com uma presente invenção.

V. Proteína semelhante ao bHLH11 (11 Hélice-Laço-Hélice básico)

Figura 20 representa a estrutura de domínio da SEQ ID N°: 245 com os motivos conservados indicados pelo sublinhado e seu número. Do domínio HLH (Motivo 8) como determinado por SMART é mostrado sublinhado em negrito.

Figura 21 representa um alinhamento múltiplo de várias proteínas semelhantes à bHLH11. Um ponto indica os resíduos conservados, um cólon indica altamente os resíduos conservados e um ponto no asterisco

para os resíduos perfeitamente conservados. O mais alto grau da conservação da sequência é observado na região do domínio bHLH. A parte do terminal C de AT2G24260 que extende-se além de outras proteínas no alinhamento é anulado.

Figura 22 O cladograma circular das proteínas bHLH selecionadas. As proteínas semelhantes à bHLH11 e uma proteína Arabidopsis que representa cada uma das outras classes definidas por Heim 2003 foram usados. O alinhamento foi gerado usando "CLUSTALX", e a árvore de ligação próxima foi calculada. O cladograma circular foi tirado usando Dendroscope (Huson et al. BMC Bioinformatics 2007). Os resultados *Bootstrap* para as 100 cópias é indicado para alguns maiores nós; o valor dos compartimentos *bootstrap* mostra que o grupo de proteínas semelhantes à bHLH11 é claramente delineado a partir de outras proteínas bHLH.

Figura 23 representa o vetor binário para a expressão aumentada em *Oryza sativa* de um ácido nucleico que codifica o bHLH11 semelhante sob o controle de um promotor GOS2 de arroz (pGOS2).

Figura 24 detalha os exemplos das sequências semelhantes à bHLH11 úteis na realização de métodos de acordo com uma presente invenção.

VI. Proteína ASR (ácido abscísico, tensão, e maturativo induzido) Figura 25 representa o vetor binário para a expressão aumentada em *Oryza sativa* de um ácido nucleico que codifica GRP sob o controle de um promotor GOS2 de arroz (pGOS2::GRP).

Figura 26 detalha os exemplos das sequências GRP úteis na realização de métodos de acordo com uma presente invenção.

VII. Proteína semelhante 11 de ligação do promotor Squamosa (SPL11)

Figura 27 representa uma sequência de aminoácido e a estrutura de domínio da SEQ ID N°: 428. Motivos conservados e os domínios são indicados. O domínio SBP é sublinhado por uma linha interrompida;

Motivo 1 é indicado em negrito; Motivo 2 é indicado em negrito e sublinhado; Motivo 3 está em letras maiúsculas em negrito e sublinhado e Motivo 4 está em negrito e sublinhado com uma linha dupla.

Figura 28 representa um alinhamento de sequência múltipla dos polipeptídeos SPL11. A posição dos resíduos de aminoácidos conservados e os domínios correspondentes aqueles descritos na Figura 27 é indicado. Uma sequência de consenso representativa dos polipeptídeos SPL11 é dado. Na sequência de consenso os aminoácidos altamente conservados são fornecidos e os espaços em branco entre estes representam qualquer aminoácido.

Figura 29 mostra uma árvore filogenética dos polipeptídeos SPL11. A árvore filogenética é como presente no Xie et al. Plant Physiology, 2006, Vol. 142, pp. 280-293, que é incoporado neste por referência como se totalmente apresentado. O grupo de polipeptídeos SPL11 no mesmo grupo como AtSPL11 (idêntico a SEQ ID N°: 2) dentro da classe nomeada Classe S3.

Figura 30 fornece uma análise da sequência dos genes miR156 de arroz (OsmiR156) e suas sequências alvos nos polipeptídeos SPL de arroz (Figura 30 A) tão bem quanto um alinhamento múltiplo dos ácidos nucleicos SPL11 (Figura 30 B). A Figura 30 A mostra um alinhamento da sequência das sequências OsmiR156 maduras com sequências complementares dos genes OsSPL. A sequência conservada de aminoácido codificada pelas sequências alvos é mostrada na parte inferior. Os pontos entre miR156 e as sequências OsSPL alvejadas indicam más combinações. A Figura 4B mostra um alinhamento múltiplo do ácido nucleico SPL11 onde altamente os resíduos de ácido nucleico conservado são indicados na sequência de consenso. A posição do local alvo miR156 altamente conservado é indicado em negrito na sequência de consenso. A SEQ ID N°: 431, SEQ ID N°: 440 e SEQ ID N°: 454 representativas dos ácidos nucleicos SPL11 que são miR156 insensíveis

não tem o local alvo miR156 conservado ou mostra uma boa divergência na sua sequência naquela posição.

Figura 31 representa o vetor binário para a expressão aumentada em *Oryza sativa* da SEQ ID N°: 427 sob o controle de um promotor GOS2 de arroz (pGOS2) (Figura 31A) ou sob um promotor WSI18 de arroz (Figura 31 B).

Figura 32 detalha os exemplos das sequências SPL11 úteis na realização de métodos de acordo com uma presente invenção.

Exemplos

A presente invenção será descrita com referência aos seguintes exemplos, que são por meio apenas da ilustração. Os seguintes exemplos não são pretendidos completamente definir ou de outra maneira limitar o escopo da invenção.

Manipulação de DNA: a não ser outra maneira estado, as técnicas de DNA recombinante são realizados de acordo com os protocolos padrões descritos em Sambrook (2001) Molecular Cloning: a laboratory manual, 3° Edition Cold Spring Harbor Laboratory Press, CSH, New York, or in Volumes 1 and 2 of Ausubel et al. (1994), Current Protocols in Molecular Biology, Current Protocols.

Materiais padrões e métodos para o trabalho molecular vegetal são descritos em Plant Molecular Biology Labfax (1993) by R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (UK) e Blackwell Scientific Publications (UK).

I. Domínio LOB que compreende a proteína (LOB: Fronteiras orgânicas laterais)

Exemplo 1: A identificação das sequências relacionadas a uma sequência do ácido nucleico usado nos métodos da invenção

As sequências (cDNA eSTs de comprimento total ou genômico) relacionadas a uma sequência do ácido nucleico usado nos

métodos da presente invenção onde identificados entre aqueles mantidos nos bancos de dados Entrez Nucleotides no National Center for Biotechnology Information (NCBI) usando ferramentas de sequência de banco de dados, tal como a ferramenta de alinhamento local básica (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa foi usado para observar as regiões da similaridade local entre as sequências em comparação com o ácido nucleico ou sequências de polipeptídeos de um banco de dados da sequência e pelo cálculo da significância estatística de pontos. Por exemplo, o polipeptídeo codificado pelo ácido nucleico usado na presente invenção foi usado pelo algorítmo TBLASTN, com ajustes default e o filtro para ignorar a baixa complexidade da sequência apresentada. A produção da análise foi examinada pela comparação em pares, e classificada de acordo com as contagens de probabilidade (valor E), onde a contagem reflete a probabilidade que um alinhamento particular ocorra pela possibilidade (o valor E inferior, quanto mais significante o acerto). Além disso aos valores E, em comparações onde também as contagens pela identidade da porcentagem. A identidade da porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácido nucleico comparado (ou polipeptídeo) em um comprimento particular.

Tabela A1 fornece uma lista das sequências de ácidos nucleicos relacionados à sequência LBD de ácido nucleico usados nos métodos da presente invenção.

Tabela A1: Exemplos de polipeptídeos LBD: sequências de DNA adicionadas de arroz, milho, trigo, canola, batata, soja, Arabidopsis

Fonte vegetal*	**SEQ ID N°. do ácido nucleico**	**SEQ ID N°. da proteína**
Arabidopsis thaliana Proteína LBD	1	2
0s01 g03890, *Oryza sativa*	59	11
0s01 g32770, *Oryza sativa*	60	12
0s03g33090, *Oryza sativa*	61	13
0s03g41330, *Oryza sativa*	62	14
0s07g40000, *Oryza sativa*	63	15
TC9404, Nicotiana benthamiana		16
TC227562, Glycine max		17
TC216138, Glycine max		18
TC147776, Hordeum vulgare		19
TC104758, Sorghum bicolor		20
TC18561, Aquilegia sp.		21
TC60668, Vitis vinifera		22
TC15459, Lotus japonicus		23
TC30552, Gossypium hirsutum		24
TC235711, Triticum aestivum	71	25
TC133081, Solanum tuberosum		26
TC107091, Medicago truncatula		27
TC147808, Hordeum vulgare		28
TC55931, Vitis vinifera		29
TC162239, Solanum tuberosum		30
TC69225, Pinus taeda		31
TC67269, Pinus taeda		32
TC220806, Glycine max		33
TC270332, Triticum aestivum		34
TC18329, Aquilegia sp.		35
TC137193, Solanum tuberosum		36
TC133385, Solanum tuberosum		37
TC140088, Solanum tuberosum		38
TC14656, Picea alba		39
TC59178, Pinus taeda		40
TC67974, Pinus taeda		41

TC178827, Lycopersicon esculentum		42
Pt-111.589, Populus tremuloides		43
Pt-V.543, Populus tremuloides		44
Pt-XIV.94, Populus tremuloides		45
Pt-11105, Populus tremuloides		46
Pt-123.86, Populus tremuloides		47
Pt-X180, Populus tremuloides		48
Pt-XII.481, Populus tremuloides		49
DQ787782, Caragana korshinskii		50
AAP37970, Brassica napus		51
ABE82505, Medicago truncatula	72	52
ABE78739, Medicago truncatula		53
Q9SN23, *Arabidopsis thaliana*	64	54
Q9SZE8, *Arabidopsis thaliana*	65	55
Q9ZW96, *Arabidopsis thaliana*	66	56
Q9M886, *Arabidopsis thaliana*	67	57
Q9CA30, *Arabidopsis thaliana*	68	58
Ls_LBD, Linum usitatissimum	69	70

*: Números de acessão dos bancos de dados GenBank ou SwissProt são fornecidos onde disponíveis, os códigos TC são a partir de TIGR TIGR.

Exemplo 2: Alinhamento das sequências LBD de polipeptídeo

O alinhamento das sequências de polipeptídeos foi realizado usando o programa AlignX a partir do vetor NTI (Invitrogen) que é baseado no algorítmo Clustal W popular do alinhamento progressivo (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31:3497-3500). Os valores default foram pela penalidade aberta de fenda de 10, pela penalidade de extensão de fenda de 0,1 e a matriz do peso selecionado é Blosum 62. A menor edição do manual foi feita ainda para otimizar o alinhamento. A conservação da sequência entre os polipeptídeos LBD foi essencialmente nos domínios dos polipeptídeos DUF260 do terminal N, a região do terminal C usualmente sendo mais variável no comprimento da sequência e composição. Os polipeptídeos LBD são alinhados na Figura 2.

A uma árvore filogenética de polipeptídeos LBD (Figura 3) foi

construído usando um algorítmo de ajustamento de ligação próxima como fornecido no programa AlignX a partir do vetor NTI (Invitrogen). As sequências das proteínas LBD de classe I usado nas construção de três são publicados disponíveis e são indicados com seus números de Acessão GenBank ou SwissProt.

Exemplo 3: Cálculo da identidade global da porcentagem entre as sequências de polipeptídeos úteis na realização dos métodos da invenção

As porcentagens globais de similaridade e identidade entre as sequências de polipeptídeos de comprimento total úteis na realização de métodos da invenção foram determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Ferramenta de alinhamento global de matriz) (BMC Bioinformatics. 2003 4:29. MatGAT: uma aplicação que gera as matrizes de similaridade/identidade usando as sequências de DNA ou proteínas. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). Software MatGAT gera as matrizes de similaridade/identidade para sequência de proteínas ou DNA sem necessariamento o pré-alinhamento dos dados. O programa realiza uma série de alinhamento em pares usando o algorítmo de alinhamento global Myers and Miller (com uma penalidade de abertura de fenda de 12, e a penalidade de extensão de fenda de 2), calcula a similaridade e identidade usando por exemplo Blosum 62 (para os polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade da sequência é mostrada na metada da parte inferior da linha de divisão e a identidade de sequência é mostrada na metada da parte superior da linha de divisão diagonal.

Parâmetros usados na comparação foram:

Matriz de contagem:	Blosum62
Primeira fenda:	12
Extensão da fenda:	2

Os resultados da análise do software são mostrados na Tabela

B pela similiaridade e identidade global no comprimento total de uma sequência de polipeptídeos. A identidade da porcentagem é dada acima em diagonal em negrito e a porcetagem da similaridade é dada abaixo em diagonal (face normal).

A identidade da porcentagem entre as sequências LBD de polipeptídeo úteis na realização de métodos da invenção podem ser tão baixo quanto 30,2 % da identidade de aminoácido comparado a SEQ ID N°: 2 (representado pela At-LBD37, linha 44). A % de identidade será igualmente mais altas onde apenas as sequências do domínio DUF206 são comparados. Para identificar o domínio DUF260 (como delineado na Figura 1) em outras proteínas LBD, o alinhamento múltiplo da Figura 2 pode ser usado.

Tabela A2: Resultados MatGAT para a similaridade e identidade global no comprimento total das sequências LBD de polipeptídeo.

	1	2	3	4	5	6	7	8	9	10	11	12
1. Os-Os01 g03890		**43,2**	**35,8**	**34,5**	**34,4**	**32,9**	**33,0**	**35,5**	**35,2**	**34,5**	**31,5**	**32,2**
2. Os-Os01 g32770	55,5		**31,0**	**31,1**	**31,5**	**30,3**	**32,4**	**30,7**	**30,6**	**30,9**	**29,8**	**31,2**
3. Os-Os03g33090	46,9	40,3		**48,3**	**65,0**	**56,7**	**53,9**	**52,1**	**66,7**	**71,3**	**51,5**	**55,0**
4.Os-0s03g41330	46,3	43,0	60,3		**48,3**	**43,0**	**43,2**	**41,2**	**46,4**	**48,6**	**40,2**	**45,1**
S.Os-0s07g40000	45,9	40,9	73,0	59,9		**53,0**	**50,6**	**49,6**	**68,0**	**63,1**	**47,6**	**52,3**
6. Nb-TC9404	45,9	42,4	66,5	60,3	67,7		**60,7**	**60,3**	**51,7**	**51,5**	**56,7**	**66,5**
7. Gm-TC227562	45,0	44,5	65,2	59,1	65,2	75,2		**73,7**	**50,6**	**49,4**	**55,2**	**64,8**
8. Gm-TC216138	47,2	41,5	63,3	58,8	62,1	70,4	79,6		**50,0**	**48,8**	**53,3**	**65,3**
9. Hv-TC147776	45,9	41,5	76,1	58,6	79,6	69,2	67,4	63,8		**58,7**	**47,3**	**52,2**
10. Sb-TC104758	45,3	40,6	77,0	59,6	71,1	63,0	62,6	58,8	68,5		**49,0**	**51,7**
11. Aq-TC18561	47,2	43,3	61,5	56,9	62,3	67,5	69,7	63,8	67,1	61,7		**58,1**
12. Vv-TC60668	45,0	43,3	65,4	59,1	66,7	79,4	77,4	77,1	66,2	64,7	72,7	
13. Lj-TC15459	45,6	40,9	64,1	59,1	64,1	72,2	81,4	88,8	67,5	61,6	70,5	79,3
14. Gh-TC30552	46,3	43,0	66,8	63,4	64,7	75,9	78,4	75,4	66,8	64,3	72,8	82,8
15. Ta-TC235711	56,7	61,5	42,9	43,2	46,3	46,7	46,3	47,4	47,0	44,6	48,4	45,6
16. St-TC133081	59,0	55,8	47,4	48,5	47,1	50,7	49,6	47,4	49,3	48,9	50,4	50,0
17. Mt-TC107091	58,6	60,0	45,5	46,2	43,5	46,5	46,2	47,2	42,5	44,9	49,2	49,2
18. Hv-TC147808	59,0	64,8	42,1	42,1	44,7	44,4	46,7	46,0	45,0	44,7	44,4	45,0
19. Vv-TC55931	57,0	55,2	49,4	47,8	46,2	53,0	51,4	49,8	48,2	46,2	53,8	51,8
20. St-TC162239	49,5	48,2	50,4	52,6	51,7	54,7	54,7	55,0	46,6	51,1	57,7	56,8
21. Pta-TC69225	53,0	50,0	38,9	41,6	40,4	44,9	44,9	43,1	43,1	41,9	44,9	44,6
22. Pta-TC67269	51,8	52,7	39,6	41,4	39,3	42,6	42,9	41,4	43,2	40,8	44,0	42,6
23. Gm-TC220806	45,9	46,1	56,4	49,1	51,3	54,5	52,6	51,7	52,3	48,5	52,4	53,5
24. Ta-TC270332	41,4	40,6	56,5	84,5	59,1	61,3	59,1	57,1	60,0	57,9	59,7	58,7
25. Aq-TC18329	62,5	61,5	46,6	47,9	44,1	50,0	50,7	47,6	44,8	45,2	51,4	49,7
26. St-TC137193	45,6	42,1	68,7	57,3	65,9	89,3	72,6	70,4	,68,9	61,3	74,0	77,6
27. St-TC133385	45,6	41,8	70,8	56,9	66,4	87,9	71,3	68,3	66,2	62,1	70,6	76,3
28. St-TC140088	44,3	42,1	64,5	51,3	61,5	71,0	68,3	65,8	65,8	60,9	68,4	70,2
29. Pa-TC14656	48,5	45,2	47,7	46,9	48,4	51,6	52,7	52,7	47,3	48,0	53,4	53,8
30. Pta-TC59178	47,7	45,2	40,5	40,2	43,5	44,1	43,3	44,6	44,1	40,5	46,8	46,6
31. Pta-TC67974	48,2	45,2	50,0	48,2	46,4	52,5	52,2	54,3	48,6	49,3	54,0	54,3
32. Le-TC178827	45,9	41,8	68,3	57,8	66,8	89,7	73,9	71,7	67,6	62,1	73,6	78,1

33. Pt-111.589	45,0	43,9	63,0	60,1	62,6	73,5	76,5	79,6	68,5	64,3	69,7	79,0
34. Pt-V.543	48,2	43,9	64,5	60,3	60,7	74,0	76,4	77,7	66,5	63,6	68,6	77,3
35. Pt-XIV.94	41,7	39,4	64,4	56,0	58,0	63,4	61,3	58,8	61,3	55,7	64,9	64,5
36. Pt-11105	43,0	39,1	63,9	54,7	58,0	60,7	63,0	58,8	65,3	56,6	63,6	65,8
37. Pt-123.86	57,7	57,9	47,6	46,3	46,3	46,6	48,6	47,6	43,6	45,9	47,3	45,9
38. Pt-X180	52,1	49,4	50,2	51,0	49,8	54,8	55,6	55,2	48,5	50,2	56,0	57,7
39. Pt-X11.481	51,5	47,3	43,6	45,8	44,7	48,4	46,5	49,5	46,5	46,9	46,9	50,9
40. Ck-LOB-	48,5	43,3	66,7	59,5	60,6	73,2	82,3	82,1	65,8	61,7	71,4	77,5
DQ787782												
41. Bn-AAP37970	56,0	56,1	48,9	49,6	46,2	49,6	50,4	51,9	50,0	47,0	50,0	51,5
42. Mt-ABE82505	47,2	44,2	65,2	62,2	65,2	72,5	79,8	80,8	64,8	64,3	71,7	76,8
43. Mt-ABE78739	54,7	56,7	45,7	48,9	44,6	50,7	50,4	49,6	47,1	47,8	51,1	50,4
44. At-LBD37-	45,0	46,1	61,6	56,4	64,0	68,8	68,8	69,6	64,0	63,6	64,0	68,8
Q9FN11												
45. At-LBD38-	45,9	45,2	61,9	55,5	63,2	61,9	70,4	70,4	62,8	63,6	67,6	70,9
Q9SN23												
46. At-LBD39-	45,0	41,8	61,3	56,7	63,3	65,8	70,8	67,9	60,0	62,5	70,4	71,7
Q9SZE8												
47. At-LBD40-	53,1	49,1	54,5	54,1	55,8	55,8	54,9	52,5	53,2	56,2	60,9	56,7
Q9ZW96												
48. At-LBD41-	52,8	56,7	44,9	49,0	47,5	49,8	50,6	47,9	49,8	47,1	51,0	50,2
Q9M886												
49. At-LBD42-	52,4	49,4	54,1	50,2	51,9	58,4	57,5	52,9	54,1	52,8	55,4	55,4
Q9CA30												

	13	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
1. Os-Os01g03890	**33,7**	**33,0**	**42,5**	**46,6**	**42,9**	**43,5**	**42,8**	**37,7**	**36,0**	**36,8**	**38,1**	**32,2**
2. Os-Os01 g32770	**30,1**	**32,4**	**52,2**	**46,5**	**47,1**	**54,4**	**41,7**	**37,0**	**37,1**	**38,2**	**39,7**	**30,9**
3. Os-Os03g33090	**54,0**	**54,5**	**32,4**	**37,0**	**34,4**	**31,3**	**37,2**	**37,6**	**30,2**	**28,3**	**41,3**	**47,4**
4. Os-Os03g41330	**41,2**	**44,1**	**32,1**	**33,1**	**30,2**	**31,4**	**36,4**	**37,0**	**29,9**	**28,4**	**35,5**	**78,7**
5. Os-Os07g40000	**50,0**	**51,7**	**34,7**	**35,8**	**33,9**	**33,2**	**35,2**	**36,5**	**28,7**	**28,3**	**38,7**	**46,3**
6. Nb-TC9404	**60,8**	**63,7**	**33,7**	**35,8**	**33,2**	**32,1**	**38,7**	**38,1**	**32,3**	**29,8**	**39,6**	**44,0**
7. G m -TC227562	**72,4**	**67,6**	**33,1**	**34,7**	**33,2**	**33,1**	**36,6**	**36,2**	**32,3**	**30,7**	**39,3**	**42,8**
8. G m -TC216138	**81,5**	**66,5**	**34,4**	**34,2**	**33,2**	**33,2**	**37,0**	**35,7**	**31,1**	**30,7**	**37,6**	**42,3**
9. Hv-TC147776	**50,8**	**51,9**	**34,9**	**35,4**	**31,6**	**32,6**	**36,8**	**34,6**	**29,3**	**28,6**	**37,9**	**45,4**
10. Sb-TC104758	**49,4**	**51,3**	**33,8**	**34,5**	**32,5**	**33,1**	**35,3**	**38,0**	**29,6**	**30,1**	**37,4**	**46,9**
11. Aq-TC18561	**54,8**	**58,5**	**33,4**	**35,0**	**34,2**	**32,1**	**37,5**	**37,6**	**32,3**	**30,8**	**37,3**	**41,1**
12. Vv-TC60668	**65,6**	**68,3**	**33,1**	**35,6**	**35,9**	**32,1**	**36,2**	**38,5**	**33,3**	**32,0**	**39,6**	**43,5**
13. Lj-TC15459		**64,9**	**34,5**	**33,2**	**33,8**	**33,4**	**36,0**	**36,5**	**34,1**	**29,4**	**36,8**	**39,6**
14. Gh-TC30552	76,8		**36,2**	**35,4**	**33,9**	**34,1**	**37,5**	**36,9**	**32,0**	**30,7**	**40,1**	**44,2**
15. Ta-TC235711	47,4	49,5		**45,5**	**49,5**	**89,9**	**49,1**	**41,3**	**37,8**	**36,4**	**46,0**	**31,8**
16. St-TC133081	50,0	50,4	57,5		**56,6**	**47,4**	**52,5**	**45,4**	**37,4**	**40,1**	**50,0**	**32,5**
17. Mt-TC107091	48,2	47,5	61,5	66,4		**50,5**	**48,8**	**41,6**	**36,4**	**39,5**	**51,7**	**31,2**
18. Hv-TC147808	45,4	47,7	92,4	59,6	63,2		**48,3**	**39,0**	**37,8**	**39,1**	**43,7**	**31,8**
19. Vv-TC55931	51,8	51,4	60,3	63,1	61,1	57,9		**47,0**	**35,8**	**38,1**	**51,2**	**36,4**
20. St-TC162239	54,9	55,1	55,7	58,8	52,8	52,3	63,2		**32,9**	**36,0**	**48,7**	**37,7**
21. Pta-TC69225	46,4	44,0	50,9	51,5	53,0	51,2	47,9	43,1		**44,5**	**32,9**	**27,8**
22. Pta-TC67269	42,3	44,3	50,9	54,8	52,4	52,4	50,0	46,4	63,1		**35,3**	**28,6**
23. Gm-TC220806	51,1	54,7	51,6	55,8	55,5	49,0	57,3	60,7	40,1	42,0		**35,8**
24. Ta-TC270332	54,9	61,2	44,3	45,3	44,2	42,7	47,8	50,9	41,0	40,2	51,7	
25. Aq-TC18329	49,7	51,0	63,8	70,0	70,8	65,2	65,9	55,5	52,1	53,6	54,1	45,9
26. St-TC137193	71,3	74,6	46,0	48,9	48,5	45,4	52,6	52,1	43,1	41,7	57,1	58,7
27. St-TC133385	70,0	73,3	43,9	48,9	46,5	43,0	52,6	53,0	41,9	41,4	57,1	57,0
28. St-TC140088	66,7	70,3	45,3	47,4	46,8	43,0	53,0	52,6	42,5	41,1	55,3	52,2
29. Pa-TC14656	51,6	53,4	50,2	49,8	48,2	50,0	46,6	52,3	47,3	46,1	44,8	44,8
30. Pta-TC59178	44,4	45,7	41,6	44,9	46,6	44,9	44,4	42,7	47,4	49,6	39,1	40,2
31. Pta-TC67974	50,7	54,3	49,1	50,0	48,2	47,0	47,8	51,8	46,4	48,2	46,7	47,8

32. Le-TC178827	72,6	75,4	46,7	48,9	48,2	43,7	53,0	52,1	42,5	43,2	56,9	58,3
33. Pt-111.589	78,6	81,9	50,5	51,5	48,5	47,0	55,3	57,1	44,3	44,6	53,4	58,0
34. Pt-V.543	76,9	79,8	49,1	52,2	49,2	48,0	53,0	56,2	45,2	45,5	53,3	57,0
35. Pt-XIV.94	59,9	62,5	44,9	45,6	42,5	44,0	48,2	51,7	41,6	41,4	56,3	57,8
36. Pt-11105	61,6	62,5	43,9	47,4	43,5	42,1	47,4	55,1	38,9	40,8	56,7	55,2
37. Pt-123.86	47,6	48,3	56,1	67,2	75,4	55,3	61,8	54,7	53,3	51,5	55,1	43,9
38. Pt-X180	56,0	56,8	53,0	56,6	55,1	51,0	66,4	59,8	48,8	47,6	60,2	49,8
39. Pt-X11.481	48,4	50,2	52,3	58,0	53,8	49,7	58,2	54,6	45,5	49,1	51,3	44,0
40. Ck-LOB-DQ787782	80,2	79,3	49,1	51,8	46,5	47,0	53,0	54,3	43,7	44,9	53,2	58,0
41. Bn-AAP37970	47,0	51,1	58,9	65,7	67,4	56,0	60,2	57,6	47,6	48,5	56,8	47,3
42. Mt-ABE82505	78,9	77,7	50,2	50,7	47,8	48,7	54,2	54,3	44,9	45,8	52,8	61,8
43. Mt-ABE78739	53,6	50,4	59,2	66,9	71,8	60,3	64,4	54,0	52,4	49,7	56,5	46,4
44. At-LBD37-Q9FN11	67,6	75,2	48,8	54,0	49,5	47,0	54,9	55,6	45,2	44,6	52,0	56,4
45. At-LBD38-Q9SN23	63,6	73,7	48,4	51,1	48,8	47,0	51,8	55,1	46,4	46,1	52,2	53,0
46. At-LBD39-Q9SZE8	73,8	70,0	48,8	50,4	47,8	47,4	52,6	53,8	45,2	43,8	51,7	61,7
47. At-LBD40-Q9ZW96	53,2	56,2	54,4	61,7	62,1	51,3	60,9	61,5	48,5	48,2	62,2	52,8
48. At-LBD41-Q9M886	49,8	52,1	56,8	67,5	67,4	57,0	61,2	58,9	49,1	49,1	57,4	46,0
49. At-LBD42-Q9CA30	56,1	54,5	54,0	54,7	54,5	51,3	60,5	64,5	47,0	45,8	62,7	51,1

	25	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**
1. Os-Os01 g03890	44,0	32,9	33,6	32,2	32,4	32,5	32,1	33,5	33,7	33,7	33,2	33,2
2. Os-Os01 g32770	46,3	30,0	29,4	29,7	28,1	29,4	30,4	30,9	30,6	29,1	30,0	31,5
3. Os-Os03g33090	33,3	61,0	60,2	50,0	35,7	33,9	38,7	60,0	52,9	52,9	50,2	51,7
4. Os-Os03g41330	31,7	44,0	44,0	39,0	36,7	30,2	36,1	44,7	41,0	41,0	41,0	39,7
5. Os-Os07g40000	34,1	52,6	52,5	46,8	35,3	33,3	37,5	51,9	49,6	48,6	46,5	46,0
6. Nb-TC9404	36,6	83,2	82,3	57,5	38,9	36,9	41,9	83,2	64,2	64,9	50,2	49,1
7. Gm-TC227562	35,2	61,1	60,3	51,0	39,8	36,6	41,1	62,1	65,3	64,5	48,5	48,5
8. Gm-TC216138	34,1	61,4	60,2	51,7	38,2	34,9	40,2	61,7	67,2	66,8	47,5	46,9
9. Hv-TC147776	32,8	54,2	54,6	50,2	33,8	32,7	35,7	53,6	50,0	49,8	46,2	47,3
10. Sb-TC104758	31,4	51,7	51,5	46,7	35,6	33,1	37,5	49,0	52,5	50,2	43,8	45,1
11. Aq-TC18561	35,5	59,3	58,8	52,8	38,3	36,4	40,5	58,0	56,8	59,4	52,6	48,5
12. Vv-TC60668	35,1	67,1	67,2	54,5	39,6	36,6	42,3	66,5	69,4	66,7	50,9	50,4
13. Lj-TC15459	33,4	59,7	59,2	51,0	37,4	37,6	39,2	60,7	65,6	64,9	47,9	47,9
14. Gh-TC30552	35,2	63,9	63,4	57,3	39,4	37,2	41,6	64,2	71,4	70,9	50,0	50,0
15. Ta-TC235711	50,3	33,8	32,8	32,4	29,5	31,0	30,8	34,5	36,2	34,8	33,4	34,5
16. St-TC133081	57,4	33,2	34,7	32,5	30,9	32,1	32,6	35,4	36,4	36,0	35,0	34,2
17. Mt-TC107091	56,8	34,2	33,6	33,1	31,5	32,0	31,0	33,9	33,8	35,2	29,6	30,8
18. Hv-TC147808	50,6	34,4	31,5	29,1	31,4	31,6	30,3	33,1	34,8	34,8	32,1	32,5
19. Vv-TC55931	56,3	37,5	39,2	36,6	33,8	31,2	34,9	38,4	37,3	35,5	36,8	35,9
20. St-TC 162239	41,8	36,3	37,2	35,0	35,6	32,0	35,7	37,6	39,0	37,3	37,6	37,9
21. Pta-TC69225	36,9	30,5	31,6	29,6	30,5	30,7	31,2	29,9	31,7	31,8	31,1	30,2
22. Pta-TC67269	43,1	30,1	29,2	29,2	28,5	30,0	32,2	31,5	29,5	31,0	29,8	29,8
23. Gm-TC220806	50,0	43,4	42,1	38,4	34,7	28,1	35,5	42,8	39,6	38,5	40,5	41,9
24. Ta-TC270332	32,1	45,5	43,5	39,2	35,7	30,3	36,8	43,3	41,3	40,9	40,9	39,7
25. Aq-TC18329		35,5	35,2	31,4	30,2	33,1	31,5	34,8	33,4	35,5	33,4	34,1
26. St-TC137193	48,6		96,8	58,4	38,8	36,1	41,9	95,0	64,3	62,5	50,2	50,2
27. St-TC133385	48,6	96,8		**59,7**	**38,1**	**35,8**	**40,1**	**92,8**	**63,9**	**60,9**	**50,7**	**50,9**
28. St-TC140088	47,9	72,4	71,9		**38,8**	**33,2**	**39,8**	**58,7**	**54,1**	**53,8**	**46,3**	**45,6**
29. Pa-TC14656	49,7	51,6	51,6	54,2		**38,6**	**88,1**	**38,7**	**42,1**	**40,3**	**33,5**	**33,5**
30. Pta-TC59178	46,0	44,9	44,6	43,3	54,3		**39,4**	**36,9**	**36,4**	**38,4**	**32,5**	**31,4**
31. Pta-TC67974	49,3	52,2	51,4	52,9	92,8	53,2		**41,7**	**42,5**	**42,8**	**35,0**	**36,1**

32. Le-TC178827	48,3	97,2	95,0	71,6	50,9	45,2	51,4		**64,6**	**62,2**	**49,5**	**49,5**
33. Pt-111.589	49,7	73,5	72,3	65,5	54,2	44,9	56,5	75,2		**83,3**	**50,4**	**48,7**
34. Pt-V.543	50,3	73,6	71,5	64,0	54,2	46,8	56,2	73,1	88,8		**50,8**	**48,3**
35. Pt-XIV.94	47,9	63,1	63,2	62,7	47,7	39,7	51,1	61,0	63,4	62,0		**86,1**
36. Pt-11105	47,2	63,1	64,6	60,4	48,4	39,4	51,8	61,5	62,2	59,9	91,8	
37. Pt-123.86	70,9	48,0	46,6	48,6	49,7	47,1	50,7	46,6	49,0	50,0	43,9	42,9
38. Pt-X180	61,0	51,9	51,9	53,5	46,2	41,6	46,7	53,5	58,5	58,3	50,2	48,5
39. Pt-X11.481	60,0	50,5	49,5	46,9	49,5	41,6	50,0	50,9	49,5	49,5	45,1	42,1
40. Ck-LOB-DQ787782	47,2	72,3	71,4	68,8	53,4	44,1	54,7	71,9	79,4	78,9	60,6	62,3
41. Bn-AAP37970	65,2	52,3	50,8	47,0	47,7	44,1	50,0	50,8	49,6	51,5	50,4	49,2
42. Mt-ABE82505	50,3	70,4	69,1	67,0	57,4	46,8	55,1	71,2	76,9	78,9	62,7	62,2
43. Mt-ABE78739	67,2	50,0	50,0	50,7	54,7	43,3	53,2	48,2	49,6	48,6	50,0	48,9
44. At-LBD37-Q9FN11	51,4	66,8	65,6	64,4	54,2	47,9	52,9	67,2	76,0	74,0	60,8	60,4
45. At-LBD38-Q9SN23	49,7	65,6	64,4	60,3	52,3	44,9	53,6	66,0	75,3	72,1	58,7	61,5
46. At-LBD39-Q9SZE8	51,0	66,3	64,6	62,1	54,5	46,8	53,3	67,9	71,7	71,1	59,6	60,8
47. At-LBD40-Q9ZW96	61,7	56,7	55,4	54,9	49,1	41,0	49,6	55,8	55,9	55,4	50,6	51,5
48. At-LBD41-Q9M886	66,9	50,6	49,8	48,7	50,5	45,5	49,3	52,1	48,7	51,3	49,0	47,1
49. At-LBD42-Q9CA30	58,3	53,6	54,1	55,4	47,7	41,3	48,2	54,5	54,6	56,2	54,9	54,5

	37	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**
1. Os-Os01 g03890	44,0	39,2	34,8	35,5	44,4	34,9	40,5	31,9	33,3	32,4
2. Os-Os01 g32770	45,3	38,5	32,3	32,1	42,7	33,9	44,8	30,2	32,1	29,7
3. Os-Os03g33090	34,7	36,9	32,7	53,0	36,4	54,9	35,1	50,0	51,2	49,6
4. Os-Os03g41330	33,7	35,2	31,2	43,6	36,2	44,6	32,6	43,9	42,4	40,7
5. Os-Os07g40000	32,8	35,7	31,9	49,8	32,6	51,3	33,1	49,2	47,4	48,1
6. Nb-TC9404	33,8	37,3	32,6	58,6	32,5	61,1	35,1	56,9	54,0	56,3
7. Gm-TC227562	32,8	36,4	32,4	76,6	34,7	74,9	37,1	56,2	57,8	57,3
8. Gm-TC216138	31,6	36,7	31,6	74,3	36,0	74,0	34,9	53,9	55,0	54,1
9. Hv-TC147776	31,1	38,2	33,5	50,6	34,3	50,6	34,5	49,2	48,6	45,8
10. Sb-TC104758	33,1	35,2	31,3	48,7	33,7	50,2	33,1	50,6	48,2	46,7
11. Aq-TC18561	33,4	36,2	32,5	56,7	36,0	57,8	36,0	51,6	55,6	57,1
12. Vv-TC60668	33,4	39,7	34,3	65,3	33,7	66,8	36,0	54,2	57,4	58,8
13. Lj-TC15459	32,1	35,9	31,0	70,4	32,8	70,7	35,3	52,8	53,0	54,8
14. Gh-TC30552	33,7	38,6	33,7	66,2	35,2	69,0	32,9	62,4	63,5	58,7
15. Ta-TC235711	45,3	42,7	36,2	35,4	46,0	38,3	46,6	34,1	35,9	33,7
16. St-TC133081	55,7	45,6	37,3	36,5	55,8	37,2	54,9	37,1	35,8	36,5
17. Mt-TC107091	63,7	43,0	37,3	34,2	54,3	33,9	60,6	33,4	37,2	34,2
18. Hv-TC147808	45,0	41,9	34,3	34,3	45,2	37,4	46,8	32,4	35,4	32,8
19. Vv-TC55931	52,0	51,4	40,9	36,8	49,8	38,7	48,9	37,4	34,9	36,7
20. St-TC162239	41,6	45,1	39,7	36,3	41,8	36,0	42,9	35,4	35,7	35,0
21. Pta-TC69225	37,6	33,4	29,4	33,4	34,7	32,3	38,0	31,7	31,1	32,5
22. Pta-TC67269	38,8	36,8	31,2	30,4	35,7	31,2	38,1	31,3	32,3	31,3
23. Gm-TC220806	48,7	48,8	40,1	39,7	48,1	37,9	51,3	36,5	38,0	38,4
24. Ta-TC270332	31,8	33,3	29,3	42,5	32,5	44,2	32,4	41,5	40,4	43,7
25. Aq-TC18329	61,1	48,3	40,6	33,4	51,2	36,6	55,9	35,1	32,8	34,4
26. St-TC137193	34,1	37,3	35,7	60,9	33,0	60,0	34,4	57,3	55,7	54,1
27. St-TC133385	34,0	37,3	36,4	60,1	34,5	59,6	33,8	56,5	54,9	53,1
28. St-TC140088	33,3	36,9	33,5	53,0	31,4	51,9	31,9	48,8	47,0	49,6
29. Pa-TC14656	32,1	33,3	33,2	41,1	29,3	41,7	32,8	39,0	40,0	39,9
30. Pta-TC59178	34,0	30,3	28,3	37,4	32,8	38,4	30,2	36,5	35,7	37,6
31. Pta-TC67974	33,2	34,2	34,1	42,2	31,9	41,9	34,2	40,9	42,6	40,4

32. Le-TC178827	32,8	38,9	36,5	60,8	34,1	60,7	33,1	56,5	55,5	55,3
33. Pt-II1.589	34,7	37,9	33,5	67,1	35,5	65,7	32,7	64,1	63,9	57,5
34. Pt-V.543	34,3	36,6	33,3	67,4	34,5	66,7	32,0	62,6	62,9	59,3
35. Pt-XIV.94	31,8	35,8	30,4	49,4	36,0	51,7	37,1	45,0	45,3	46,5
36. Pt-11105	32,4	35,0	30,4	48,5	36,7	51,1	35,6	43,6	45,7	48,1
37. Pt-123.86		45,8	39,3	33,1	56,2	34,3	55,9	34,8	34,1	33,0
38. Pt-X180	56,4		43,1	34,6	42,3	34,9	44,1	37,2	37,3	39,9
39. Pt-X11.481	55,4	56,8		32,5	37,5	33,6	37,4	31,9	30,5	29,9
40. Ck-LOB-DQ787782	49,0	53,9	47,3		35,5	80,9	33,8	54,0	55,8	57,8
41. Bn-AAP37970	70,3	56,8	55,7	51,5		37,7	51,6	33,2	34,1	34,3
42. Mt-ABE82505	47,0	55,2	47,3	87,1	51,9		36,7	55,6	57,3	56,7
43. Mt-ABE78739	67,6	57,9	55,4	50,7	64,7	51,4		31,9	32,7	34,3
44. At-LBD37-Q9FN11	53,0	53,2	48,0	69,2	50,4	66,8	50,0		76,9	55,3
45. At-LBD38-Q9SN23	47,0	55,9	49,5	70,0	51,5	69,2	48,6	86,4		56,8
46. At-LBD39-Q9SZE8	49,0	57,3	49,1	72,5	53,8	71,3	51,1	67,2	68,8	
47. At-LBD40-Q9ZW96	65,5	62,7	52,0	55,8	68,9	56,2	66,2	55,2	55,5	56,3
48. At-LBD41-Q9M886	70,9	58,6	57,1	47,9	90,5	50,2	65,5	53,2	51,7	52,1
49. At-LBD42-Q9CA30	55,4	62,7	52,0	58,4	56,8	55,8	57,2	53,6	52,6	52,9

	47	**48**	**49**
1. Os-Os01g03890	**41,5**	**41,3**	**38,2**
2. Os-Os01 g32770	**41,5**	**44,8**	**39,0**
3. Os-Os03g33090	**40,4**	**35,0**	**40,8**
4.Os-0s03g41330	**38,2**	**33,7**	**36,0**
S.Os-0s07g40000	**40,2**	**33,8**	**36,3**
6. Nb-TC9404	**39,4**	**34,6**	**40,7**
7. Gm-TC227562	**40,8**	**35,7**	**39,8**
8. Gm-TC216138	**36,5**	**34,0**	**35,5**
9. Hv-TC147776	**38,2**	**34,6**	**37,7**
10. Sb-TC104758	**36,2**	**34,5**	**38,6**
11. Aq-TC18561	**42,1**	**37,2**	**36,6**
12. Vv-TC60668	**40,2**	**36,1**	**38,2**
13. L j-TC 15459	**38,1**	**35,1**	**36,4**
14. Gh-TC30552	**39,8**	**34,6**	**36,6**
15. Ta-TC235711	**44,3**	**44,3**	**41,0**
16. St-TC133081	**50,7**	**55,9**	**45,7**
17. Mt-TC 107091	**52,6**	**56,1**	**40,2**
18. Hv-TC147808	**42,1**	**46,5**	**39,7**
19. Vv-TC55931	**49,2**	**49,3**	**47,3**
20. St-TC162239	**46,5**	**44,5**	**47,5**
21. Pta-TC69225	**36,4**	**36,2**	**34,1**
22. Pta-TC67269	**36,0**	**36,3**	**33,3**
23. Gm-TC220806	**52,8**	**50,8**	**49,8**
24. Ta-TC270332	**39,2**	**30,4**	**36,3**
25. Aq-TC18329	**49,8**	**53,2**	**43,7**
26. St-TC137193	**39,2**	**34,2**	**38,6**
27. St-T C 133385	**40,1**	**34,1**	**38,4**
28. St-TC140088	**37,8**	**35,7**	**36,6**

29. Pa-TC14656	**33,8**	**30,4**	**31,3**
30. Pta-TC59178	**30,2**	**32,5**	**27,2**
31. Pta-TC67974	**34,8**	**31,8**	**33,0**
32. Le-TC178827	**39,0**	**34,6**	**38,5**
33. Pt-111.589	**36,6**	**35,0**	**35,3**
34. Pt-V.543	**38,2**	**35,4**	**39,2**
35. Pt-XIV.94	**37,8**	**33,8**	**38,3**
36. Pt-11105	**40,2**	**35,7**	**37,4**
37. Pt-123.86	**53,8**	**59,1**	**43,6**
38. Pt-X180	**46,8**	**41,5**	**49,6**
39. Pt-X11.481	**38,5**	**37,6**	**36,4**
40. Ck-LOB-DQ787782	**40,2**	**31,6**	**38,2**
41. Bn-AAP37970	**55,5**	**83,5**	**42,1**
42. Mt-ABE82505	**40,3**	**33,8**	**36,8**
43. Mt-ABE78739	**53,3**	**51,6**	**43,8**
44. At-LBD37-Q9FN11	**37,5**	**34,5**	**34,9**
45. At-LBD38-Q9SN23	**40,5**	**33,5**	**34,8**
46. At-LBD39-Q9SZE8	**38,3**	**35,0**	**37,1**
47. At-LBD40-Q9ZW96		**56,7**	**46,5**
48. At-LBD41-Q9M886	68,4		**42,6**
49. At-LBD42-Q9CA30	61,8	56,3	

Exemplo 4: A identificação dos domínios compreendidos nas sequências de polipeptídeos úteis na realização de métodos da invenção

O recurso integrado das famílias de proteínas, bancos de dados de locais e domínios (InterPro) é uma interface integrada para os bancos de dados assinados comumente usados para as buscas com base em texto e sequências. O banco de dado InterPro combina estes bancos de dados, que usam metodologias diferentes e variam os graus de informação biológica sobre proteínas bem caracterizadas para derivar os sinais da proteína. A colaboração dos bancos de dados incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. O Pfam é uma ampla coleção de alinhamentos de sequência múltipla e modelos Markov secretos que cobrem muitos dos domínios de proteína comuns e famílias. O Pfam é hospedado no servidor Sanger Institute in the United Kingdom. Interpro é hospedado no European Bioinformatics Institute in the United Kingdom.

Os resultados da avaliação InterPro de uma sequência de polipeptídeo como representado pela SEQ ID N°: 2 são apresentados na Tabela A3.

Tabela A3: Resultados de avaliação InterPro (maiores números de acessão) de uma sequência de polipeptídeo como representado pela SEQ ID N°: 2.

Bancos de dados	Número de acessão	Nome de acessão	Coordenadas de ácido nucleico na SEQ ID N°
PFAM	PF03195	DUF260	2-107
PROFILE	PS50891	LOB	1-107

Exemplo 5: Predição da topologia de uma sequência de polipeptídeos úteis na realização de métodos da invenção

O TargetP 1.1 prediz a localização subcelular das proteínas eucarióticas. A indicação da localização é com base na presença predita de qualquer uma das pré-sequências do terminal N: peptídeo de transição de cloroplasto (cTP), peptídeo de alvejamento mitocrondrial (mTP) ou peptídeo de sinal de caminho secretório (SP). As contagens em que a predição final não são baseados nas reais probabilidades, e estes não são necessariamente adicionado a um. Entretanto, a localização com a contagem mais alta é igualmente de acordo com o TargetP, e a conexão entre as contagens (a classe de confiabilidade) pode ser uma indicação de que certas predições são. A clsse de confiabilidade (RC) varia de 1 a 5, onde 1 indica a predição mais forte. TargetP é mantido no servidor da Technical University of Denmark.

Para as sequências preditas por conter uma pré-sequência de terminal N e um local de clivagem potencial também pode ser predito.

Um número de parâmetros foram selecionados, tal como grupo de organismos (vegetal ou não vegetal), séries de cortes (nenhum, série pré-definida de cortes, ou série especificadas pelo uso de cortes), e o cálculo da predição dos locais de clivagem (sim ou não).

Os resultados da análise TargetP 1.1 de uma sequência de polipeptídeo como representado pela SEQ ID N°: 2 são apresentados na Tabela A4. O grupo de organismos "vegetal" foi selecionado, nos cortes definidos, e o comprimento predito do peptídeo de transição requerido.

Tabela A4: Análise TargetP 1.1 de uma sequência de polipeptídeo como

representado pela SEQ ID N°: 2

Comprimento (AA)	250
Peptídeo de transição de cloroplástico	0,026
Peptídeo de transição mitocondrial	0,401
Peptídeo de sinal do caminho secretório	0,019
Outro alvejamento subcelular	0,573
Localização predita	/
Classe de confiabilidade	5
Comprimento de peptídeo de transição predita	/

A localização subcelular de uma sequência de polipeptídeo como representado pela SEQ ID N°: 2 pode ser o citoplasma ou núcleo, nenhum peptídeo de transição predito. SubLoc (Hua & Sun, Bioinformatics 17, 721 - 728, 2001) prediz uma localização nuclear (índice de confiabilidade: 2, precisão: 74 %); esta predição está em entendimento com os dados de Liu et al (2005).

Muitos outros algorítmos podem ser usados para realizar tais análises, incluindo:

- ChloroP 1.1 hospedado no servidor da Technical University of Denmark;
- Protein Prowler Subcellular Localisation Predictor versão 1.2 hospedado no servidor do Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia;
- PENCE Proteome Analyst PA-GOSUB 2.5 hospedado no servidor da University of Alberta edmonton, Alberta, Canada;
- TMHMM, hospedado no servidor da Technical University of Denmark

Exemplo 6: Clonagem da sequência LBD de ácido nucleico usado nos métodos da invenção

A sequência do ácido nucleico usado nos métodos da invenção foi amplificado por PCR usando como modelo uma biblioteca de cDNA de mudas feitas comumente de *Arabidopsis thaliana* (em pCMV Sport 6.0; Invitrogen, Paisley, UK). O PCR foi realizado usando polimerase Hifi Taq

DNA em condições padrões, usando 200 ng do modelo em uma mistura de PCR a 50 µl. Os iniciadores usados foram prm009067 (SEQ ID N°: 3; sentido, códon de partida em negrito):

5'-ggggacaagtttgtacaaaaaagcaggcttaaacaatgagctgcaatggttgc-3' e prm009068 (SEQ ID N°: 4; reverso, complementar):

5'-ggggaccactttgtacaagaaagctgggtactaactctgagaaaaccgcc-3',

que incluem os locais AttB para a recombinação Gateway. O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante o qual fragmento PCR recombina in vivo com o plasmídio pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada", pLBD. O plasmídio pDONR201 foi adquirido de Invitrogen, como partes da tecnologia Gateway®.

O clone de entrada que compreende SEQ ID N°: 1 foi então usado na reação LR com um vetor de destino usado para a transformação *Oryza sativa*. Este vetor contido como elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; um cassete de expressão marcador avaliável; e um cassete Gateway pretendido pela recombinação LR *in vivo* com uma sequência do ácido nucleico de interesse já clonado no clone de entrada. Um promotor GOS2 de arroz (SEQ ID N°: 10) para a expressão específica da raiz foi localizada a montante deste cassete Gateway.

Após a etapa de recombinação LR, o vetor de expressão resultante pGOS2::LBD (Figura 4) foi transformado na cepa *Agrobacterium* LBA4044 de acordo com os métodos bem conhecidos na técnica.

Exemplo 7: Transformação da planta

Transformação de arroz

A *Agrobacterium* contendo o vetor de expressão foi usado para transformar as plantas *Oryza sativa*. As sementes secas maduras de japonica

cultiva de arroz Nipponbare foram descascados. A esterilização foi realizada pela incubação em um minuto em 70 % de etanol, seguido por 30 minutos em 0,2 % de $HgCl_2$, seguido por 6 vezes de 15 minutos lavado com água destilada estéril. As sementes estéreis foram então germinadas em um meio contendo 2,4-D (meio de indução de calos). Após a incubação no escuro por quatro semanas calos derivados de escutelo embriogênico foram taxados e propagados no mesmo meio. Após duas semanas, os calos foram multiplicados ou propagados pela subcultura no mesmo meio por outras 2 semanas. Os pedaços de calos embriogênicos foram sub-culturados em meio fresco 3 dias antes da co-cultivação (para intensificar a atividade de divisão celular).

A cepa da *Agrobacterium* LBA4404 contendo o vetor de expressão foi usado para a co-cultivação. A *Agrobacterium* foi inoculado no meio AB com os antibióticos apropriados e cultivados por 3 dias a 28° C. As bactérias foram então coletadas e submetida à suspensão no meio de co-cultivação líquido a uma densidade (OD_{600}) de cerca de 1. A suspensão foi então transferida a um disco Petri e os calos sumetidos à imersão na suspensão por 15 minutos. Os tecidos dos calos foram então manchados secos em um papel de filtro e transferidos para solidificar-se, o meio de co-cultivação e incubados por 3 dias no escuro a 25° C. Os calos co-cultivados foram desenvolvidos em meio contendo 2,4-D por 4 semanas no escuro a 28° C na presença de um agente de seleção. Durante este período, os calos resistentes rapidamente ao desenvolvimento foram ilhas desenvolvidas. Após a transferência deste material material a um meio de regeneração e incubação na luz, o potencial embriogênico foi liberado e os brotos desenvolvidos nas próximas quatro ou cinco semanas. Os brotos foram taxados dos calos e incubados por 2 a 3 semanas em um meio contendo auxina de que estes foram transferidos ao solo. Os brotos endurecidos foram desenvolvidos sob alta umidade e dias curtos em uma estufa para plantas.

Aproximadamente 35 transformantes de arroz TO independentes foram gerados por uma construção. Os primeiros transformantes foram transferidos a partir de uma câmera de cultura de tecido de uma estufa para plantas. Após uma análise PCR quantitativa para verificar o número de cópias da inserção de T-DNA, apenas uma simples cópia de planta transgênica que exibe a tolerância ao agente de seleção foi mantida para a coleta da semente T1. As sementes foram então coletadas três a cinco meses após a transplantação. O método produzido de transformantes de local simples em uma taxa acima de 50 % (Aldemita and Hodges1996, Chan et al. 1993, Hiei et al. 1994).

Transformação do milho

Transformação do milho *(Zea mays)* é realizado com uma modificação do método descrito por Ishida et al. (1996) Nature Biotech 14(6): 745-50. A transformação é dependente do genotipo em milho e apenas os genotipos específicos são responsáveis pela transformação e regeneração. A linha inata A188 (University of Minnesota) ou híbridos com A188 como um padrão são boas fontes de material doador para a transformação, mas outros genotipos podem ser usados sucessivamente também. As espigas são coletadas a partir da planta do milho aproximadamente 11 dias após a polinização (DAP) quando o comprimento do embrião imaturo é de cerca de 1 a 1,2 mm. Os embriões imaturos são co-cultivados com *Agrobacterium tumefaciens* contendo o vetor de expressão, e as plantas transgênicas são recuperadas através de organogênese. Os embriões taxados são desenvolvidos no meio de indução de calos, então o meio de regeneração do milho, contendo o agente de seleção (por exemplo imidazolinona mas vários marcadores de seleção podem ser usados). As placas de Petri são incubadas na luz a 25° C por 2 a 3 semanas, ou até o desenvolvimento dos brotos. Os brotos verdes são transferidos de cada embrião ao meio enraizado do milho e incubado a 25° C por 2 a 3 semanas, até o desenvolvimento das raízes. Os brotos enraizados são

transplantados no solo na estufa para plantas. As sementes T1 são produzidas a partir de plantas que exibem a tolerância ao agente de seleção e que contém uma cópia simples da inserção T-DNA.

Transformação do trigo

A transformação de trigo é realizada com o método descrito por Ishida et al. (1996) Nature Biotech 14(6): 745-50. O cultivo Bobwhite (disponível de CIMMYT, Mexico) é comumente usado na transformação. os embriões imaturos são co-cultivados com *Agrobacterium tumefaciens* contendo o vetor de expressão, e as plantas transgênicas são recuperadas através de organogênese. Após a incubação com *Agrobacterium,* os embriões são desenvolvidos in vitro em meio de indução de calos, então o meio de regeneração, contendo o agente de seleção (por exemplo imidazolinona mas vários marcadores de seleção podem ser usados). As placas de Petri são incubadas na luz a 25° C por 2 a 3 semanas, ou até o desenvolvimento dos brotos. Os brotos verdes são transferidos de cada embrião ao meio enraizado e incubado a 25° C por 2 a 3 semanas, até o desenvolvimento das raízes. Os brotos enraizados são transplantados no solo na estufa para plantas. As sementes T1 são produzidas a partir de plantas que exibem a tolerância ao agente de seleção e que contém uma cópia simples da inserção T-DNA.

Transformação da soja

A soja é transformada de acordo com uma modificação do método descrito no Texas A&M Patente U.S. 5.164.310. Diversas variedades de soja comerciais são responsáveis pela transformação deste método. O cultivo Jack (disponível de Illinois Seed foundation) é comumente usado para a transformação. As sementes de soja são esterilizadas por semeadura *in vitro*. A hipocotila, a radícula e uma cotiledônia são taxadas a partir das mudas do sétimo dia de idade. A epicotila e a cotiledônia remanescente ainda são desenvolvidas para desenvolver os nós auxiliares. Estes nós auxiliares são taxados e incubados com *Agrobacterium tumefaciens* contendo o vetor de

expressão. Após o tratamento da co-cultivação, os explantes são lavados e transferidos para o meio de seleção. Os brotos regenerados são taxados e colocados em um meio de alongamento de brotos. Os brotos não mais longos do que 1 cm são colocados em um meio enraizado até o desenvolvimento das raízes. Os brotos enraizados são transplantados no solo na estufa para plantas. As sementes T1 são produzidas a partir de plantas que exibem a tolerância ao agente de seleção e que contém uma cópia simples da inserção T-DNA.

Transformação de colza/canola

Os pecíolos cotiledônios e hipocotilas das mudas de 5 a 6 dias de idade são usadas como explantes para a cultura de tecido e transformada de acordo com o Babic et al. (1998, Plant Cell Rep 17: 183-188). O cultivo comercial Westar (Agriculture Canada) é a variedade padrão usada para a transformação, mas outras variedades também podem ser usadas. As sementes de canola são esterilizadas pela superfície para a semeadura in vitro. Os explantes de pecíolo cotiledônia com uma cotiledônia ligada são taxados a partir da semadura in vitro, e inoculado com *Agrobacterium* (contendo o vetor de expressão) pela imersão da expremidade do corte do explante de pecíolo na suspensão bacteriana. Os explantes são então cultivados por 2 dias em meio MSBAP-3 contendo 3 mg/l de BAP, 3 % de sacarose, 0,7% de Fitagar a 23° C, 16 horas de luz. Após dois dias de co-cultivo com *Agrobacterium,* os explantes de pecíolo são transferidos ao meio MSBAP-3 contendo 3 mg/l de BAP, cefotaxima, carbenicilina, ou timentina (300 mg/l) por 7 dias, e então cultivado em meio MSBAP-3 com cefotaxima, carbenicilina, ou timentina e agente de seleção até a regeneração dos brotos. Quandos os brotos são 5 a 10 mm em comprimento, estes são cortados e transferidos ao meio de alongamento de brotos (MSBAP-0,5, contendo 0,5 mg/l de BAP). Os brotos de cerca de 2 cm em comprimento são transferidos ao meio de enraizamento (MSO) para a introdução da raiz. Os brotos enraizados são transplantados no solo na estufa para plantas. As sementes T1 são produzidas a partir de plantas

que exibem a tolerância ao agente de seleção e que contém uma cópia simples da inserção T-DNA.

Transformação da Alfalfa

Um clone de regeneração da alfafa *(Medicago sativa)* é transformada usando o método de (McKersie et al., 1999 Plant Physiol 119: 839-847). A regeneração e transformação da alfafa é o genotipo independente e portanto uma regeneração da planta é requerida. Os métodos para obter a regeneração das plantas foram descritos. Por exemplo, estes podem ser selecionados a partir do cultivo Rangelander (Agriculture Canada) ou qualquer outra variedade de alfafa comercial como descrito por Brown DCW e A Atanassov (1985. Plant Cell Tissue Organ Culture 4: 111-112). Alternativamente, a variedade RA3 (University of Wisconsin) foi selecionada para o uso na cultura do tecido (Walker et al., 1978 Am J Bot 65:654-659). Os explantes de pecíolo são co-cultivados com uma cultura durante a noite de *Agrobacterium tumefaciens* C58C1 pMP90 (McKersie et al., 1999 Plant Physiol 119: 839-847) ou LBA4404 contendo o vetor de expressão. Os explantes são co-cultivados por 3 dias no escuro em meio de indução SH contendo 288 mg/L de Pro, 53 mg/ L de tioprolina, 4,35 g/L de K2SO4, e 100 µm de acetosiringinona. Os explantes são lavados em meio Murashige-Skoog com metade do comprimento (Murashige and Skoog, 1962) e colocado no mesmo meio de indução SH sem acetosiringinona mas com um agente de seleção adequada e antibiótico adequado para inibir o desenvolvimento da *Agrobacterium*. Após diversas semanas, os embriões somáticos são transferidos ao meio de desenvolvimento BOi2Y contendo nenhum regulador de desenvolvimento, nenhum antibiótico, e 50 g/ L de sacarose. Os embriões somáticos são subsequentemente germinados no meio Murashige-Skoog de comprimento médio. As mudas enraizadas foram trasnplantadas em potes e o desenvolvimento em uma estufa para plantas. As sementes T1 são produzidas a partir de plantas que exibe a tolerância ao agente de seleção e que contém

uma cópia simples da inserção T-DNA.

Transformação do algodão

A transformação do algodão *(Gossypium hirsutum* L.) é realizado usando *Agrobacterium tumefaciens,* em explantes hipocotilas. O cultivo comercial tal como Coker 130 ou Coker 312 (SeedCo, Lubbock, TX) são variedades padrões usadas para a transformação, mas outras variedades também podem ser usadas. As sementes tem a superfície esterilizada e germinadas no escuro. Os explantes de hipocotiledônia são cortados a partir das mudas germinadas aos comprimentos de cerca de 1 a 1,5 centímetro. O explante de hipocotila é submetido à imersão no inóculo de *Agrobacterium tumefaciens* contendo o vetor de expressão, por 5 minutos então co-cultivados a cerca de 48 horas em MS +1,8 mg/l de KNO3 + 2 % de glicose a 24° C, no escuro. Os explantes são transferidos o mesmo meio contendo bactéria apropriada e planta dos marcadores selecionáveis (diversas vezes repetido), até os calos embriogênicos ser visto. Os calos são separados e subcultivados até os embriões somáticos aparecerem. As plantinhas derivadas dos embriões somáticos são maduras em meio de enraizamento até o desenvolvimento das raízes. Os brotos enraizados são transportados ao solo em potes nas estufas. As sementes T1 são produzidas a partir de plantas que exibem a tolerância ao agente de seleção e que contém uma cópia simples da inserção T-DNA.

Exemplo 8: Procedimento de avaliação fenotípica

8.1 Apresentação da avaliação

Aproximadamente 35 transformantes de arroz TO independentes foram gerados. Os primeiros transformantes foram transferidos a partir de uma câmera de cultura de tecido de uma estufa para plantas para o desenvolvimento e coletado da semente T1. Seis eventos, de que a progênie T1 segregada 3:1 para a presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 mudas T1 contendo o transgene (hetero e homozigotos) e aproximadamente 10 mudas T1 perderam o

transgene (nulizigotos) foram selecionados pela expressão marcadora visual do monitoramento. As plantas transgênicas e os nulizigotos correspondentes foram desenvolvidos lado a lado em posições aleatórias. As condições da estufa foram de dias curtos (12 horas de luz), 28° C na luz e 22° C no escuro, e uma umidade relativa de 70 %. As plantas em desenvolvimento sob as condições de não tensão são fornecidos com água em intervalos regulares para garantir que a água e nutrientes não são limitantes para satisfazer as necessidades da planta para completar o crescimento e desenvolvimento.

Avaliação da eficiência do uso de nitrogênio

As plantas de arroz das sementes T2 foram desenvolvidas no solo em potes sob condições normais exceto para a solução do nutriente. Os potes foram submetidos à água a partir da transplantação do amadurecimento com uma solução de nutriente específico contendo teor de nitrogênio N reduzido (N), usualmente entre 7 a 8 vezes menor. O resto do cultivo (amadurecimento da planta, coleta da semente) foi a mesma como para as plantas não desenvolvidas sob a tensão abiótica. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais.

8.2 Análise estatística: Teste F

Dois fatores ANOVA (análise de variantes) foi usado como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformados com o gene da presente invenção. O teste F foi realizado para checar por um efeito do gene em todos os eventos de transformação e para verificar por um total do efeito do gene, também conhecido como um efeito do gene global. O limiar para a significância por um efeito verdadeiro do gene global foi apresentado em uma probabilidade de 5 % do nível para o teste F. Um ponto de valor do teste F significante por um efeito do gene, significando que este não é apenas uma mera presença ou

posição do gene que está causando as diferenças no fenotipo.

8.3 Parâmetros medidos

Medição do parâmetro relacionado à biomassa

A partir do estágio de semeadura até o estágio de maturidade das plantas foram passados diversas vezes através de um gabinete de imagem digital. Em cada período de tempo as imagens digitais (2048x1536 píxeis, 16 milhões de cores) foram tiradas de cada planta em pelo menos 6 ângulos diferentes.

Uma área da planta acima do solo (ou biomassa da folha) foi determinado pela contagem do número total de píxeis nas imagens digitais a partir das partes de plantas acima do solo discriminadas a partir do fundamento. Este valor foi medido pelas figuras retiradas do mesmo ponto de tempo a partir de ângulos diferentes e foi convertido a um valor de superfície física expressada em quadrado mm pela calibração. Os experimentos mostram que à área da planta acima do solo medida desta maneira correlaciona com a biomassa das partes das plantas trituradas acima. À área triturada acima é a área medida no ponto de tempo em que a planta tem atingido esta biomassa máxima da folha. O vigor precoce foi determinado pela contagem do número total de píxeis a partir das partes de plantas acima do solo discriminadas a partir do fundamento. Este valor foi medido pelos desenhos retirados do mesmo ponto de tempo de ângulos diferentes e foi convertido a um valor de superfície física expressada em quadrado mm pela calibração. Os resultados descritos abaixo são para as plantas com três semanas pós-germinação.

Medições dos parâmetros relacionados à semente

As panículas primárias maduras foram coletadas, contadas, ensacadas, rotuladas com código de barra e então secadas por três dias em um forno a 37° C. As panículas foram então limiadas e todas as sementes foram coletadas e contadas. As cascas enchidas foram separadas a partir de um vazio usando um dispositivo de sopro de ar. As cascas vazias foram descartadas e a

fração remanescente foi contada novamente. As cascas enchidas foram pesadas em uma balança analítica. O número de sementes enchidas foi determinado pela contagem do número de cascas enchidas que permanece depois da etapa de separação. A produção total de semente foi medida pela pesagem de todas as cascas enchidas coletadas a partir da planta. O número total de semente por planta foi medido pela contagem do número de cascas coletadas a partir da planta. Peso de núcleo milhar (TKW) é extrapolado a partir do número de sementes enchidas contadas e seu peso total. O índice de colheita (HI) na presente invenção é definida como a razão entre a produção total de semente e a área triturada acima (mm^2), multiplicada por um fator 10^6. A taxa de abastecimento da semente como definido na presente invenção é a proporção (expressado como uma %) do número de sementes enchidas no total do número de sementes (ou florzinhas).

Exemplo 9: Resultados da avaliação fenotípica das plantas transgênicas que compreendem a SEQ ID N°: 1

A avaliação das plantas transgênicas de arroz que expressa um ácido nucleico *LBD* sob as condições de não tensão mostrou que aqui foi um aumento maior do que 5 % para a biomassa acima do solo (AreaMax), produção total de semente, número de sementes enchidas, taxa de abastecimento, índice de colheita, e mais do que 3 % por peso de núcleo milhar.

A avaliação das plantas transgênicas de arroz que expressam um ácido nucleico *LBD* na avaliação da eficiência do uso de nitrogênio revelado um aumento maior do que 5 % para o vigor de emergência (vigor precoce).

Exemplo 10: Resultados da avaliação fenotípica das plantas transgênicas que compreendem a SEQ ID N°: 71

A região codificada compreendida na SEQ ID N°: 71 foi clonada sob o controle do promotor GOS2 de arroz no vetor de trasnformação

do arroz como descrito no Exemplo 6. As plantas transgênicas de arroz que compreendem a região codificadora da SEQ ID N°: 71 foram gerados seguindo os procedimentos do Exemplo 7. As plantas foram avaliadas de acordo com os procedimentos descritos no Exemplo 8. SEQ ID N°: 71 codifica a proteína LBD representada pela SEQ ID N°: 25.

A avaliação das plantas transgênicas de arroz que expressam a SEQ ID N°: 71 sob as condições de não tensão mostrou que aqui foi um aumento maior do que 5 % para a biomassa acima do solo (AreaMax), o número de flores por panícula, o número total de sementes por planta, e mais do que 3 % por peso de núcleo milhar.

Exemplo 11: Resultados da avaliação fenotípica das plantas transgênicas que compreendem a SEQ ID N°: 72

A região codificada compreendida na SEQ ID N°: 72 foi clonada sob o controle do promotor GOS2 de arroz no vetor de trasnformação do arroz como descrito no Exemplo 6. As plantas transgênicas de arroz que compreendem uma região codificadora da SEQ ID N°: 72 foram geradas seguindo os procedimentos do Exemplo 7. As plantas foram avaliadas de acordo com os procedimentos descritos no Exemplo 8. A SEQ ID N°: 72 codifica uma proteína LBD representada pela SEQ ID N°: 52.

A avaliação das plantas transgênicas de arroz que expressam a SEQ ID N°:.72 sob as condições de não tensão mostrou que aqui foi um aumento maior do que 5 % para a biomassa acima do solo (AreaMax), peso da semente por planta, número de sementes enchidas, o número de flores por panícula, o número total de sementes por planta, índice de colheita e mais do que 3 % por peso de núcleo milhar.

A avaliação das plantas transgênicas de arroz que expressam a SEQ ID N°: 72 na avaliação da eficiência do uso de nitrogênio revelado um aumento maior do que 5 % para a biomassa acima do solo (AreaMax).

II. Polipeptídeo JMJC (JUMONJI-C)

Exemplo 12: A identificação das sequências relacionadas à sequência JMJC de ácido nucleico usado nos métodos da invenção

As sequências (cDNA eSTs de comprimento total ou genômico) relacionadas a uma sequência do ácido nucleico usado nos métodos da presente invenção onde identificados entre aqueles mantidos no bancos de dados Entrez Nucleotides no National Center for Biotechnology Information (NCBI) usando ferramentas de sequência de banco de dados, tal como a ferramenta de alinhamento local básica (BLAST) (Altschul et al. (1990) J. Mot. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa é usado para observar as regiões da similaridade local entre sequências em comparação ácido nucleico ou sequências de polipeptídeos de um banco de dados da sequência e pelo cálculo da significância estatística de pontos. Por exemplo, o polipeptídeo codificado pelo ácido nucleico usado na presente invenção foi usado pelo algorítmo TBLASTN, com ajustes default e o filtro para ignorar baixa complexidade sequência apresentada. A produção da análise foi examinada pela comparação em pares, e classificada de acordo com as contagens de probabilidade (valor E), onde a contagem reflete uma probabilidade que um alinhamento particular ocorra pela possibilidade (o valor E inferior, quanto mais significante o acerto). Além disso, os valores E, em comparação também foram contados pela identidade de porcentagem. A identidade da porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácido nucleico comparado (ou polipeptídeo) em um comprimento particular.

Tabela B1 fornece uma lista das sequências de ácidos nucleicos relacionados à sequência JMJC de ácido nucleico usado nos métodos da presente invenção. Os polipeptídeos com uma acessão estende-se por um ponto e um dígito que representa as variantes de união.

Tabela B1: Exemplos de polipeptídeos JMJC

fonte vegetal	número de acessão Genbank (ou local)	SEQ ID N° do ácido nucleico	SEQ ID N° da proteína
Arabidopsis thaliana	AT3G20810	73	74
Arabidopsis thaliana	AT3G20810	83	84
Arabidopsis thaliana	AT5G19840	85	86
Arabidopsis thaliana	AT3G45880	87	88
Medicago truncatula	ABE92082	89	90
Brachypodium sylvaticum	CAJ26373	91	92
Oryza sativa	0s09g0483600	93	94
Arabidopsis thaliana	AT1G08620	95	96
Arabidopsis thaliana	AT1 G09060	97	98
Arabidopsis thaliana	AT1 G09060	99	100
Arabidopsis thaliana	AT1 G09060	101	102
Arabidopsis thaliana	AT 1 G 11950	103	104
Arabidopsis thaliana	AT1 G30810	105	106
Arabidopsis thaliana	AT1G62310	107	108
Arabidopsis thaliana	AT1G63490	109	110
Arabidopsis thaliana	AT1G78280	111	112
Arabidopsis thaliana	AT2G34880	113	114
Arabidopsis thaliana	AT2G38950	115	116
Arabidopsis thaliana	AT3G07610	117	118
Arabidopsis thaliana	AT3G48430	119	120
Arabidopsis thaliana	AT4G00990	121	122
Arabidopsis thaliana	AT4G20400	123	124
Arabidopsis thaliana	AT4G20400	125	126
Arabidopsis thaliana	AT5G04240	127	128
Arabidopsis thaliana	AT5G06550	129	130
Arabidopsis thaliana	AT5G46910	131	132
Arabidopsis thaliana	AT5G63080	133	134
Oryza sativa	0s01 g36630	135	136
Oryza sativa	0s01 g67970	137	138
Oryza sativa	0s02g01940	139	140
Oryza sativa	0s02g58210	141	142
Oryza sativa	0s02g58210	143	144
Oryza sativa	0s03g05680	145	146
Otyza sativa	0s03g22540	147	148
Oryza sativa	0s03g27250	149	150
Oryza sativa	0s03g31594	151	152
Oryza sativa	0s03g31594	153	154
Oryza sativa	0s05g 10770	155	156
Oryza sativa	0s05g23670	157	158

Otyza sativa	0s09g22540	159	160
Oryza sativa	0s10g42690	161	162
Oryza sativa	0s11g36450	163	164
Oryza sativa	Os 12g 18149	165	166
Oryza sativa	Os 12g 18150	167	168
Glycine max	Gm_JMJ_1	169	170

Em alguns exemplos, as sequências relacionadas tem sido como tentativa reunido e publicado descoberto pelas instituições de busca, tal como The Institute for Genomic Research (TIGR). Os bancos de dados Eukaryotic Gene Orthologs (EGO) podem ser usados para identificar tais sequência relacionadas pela busca de palavra chave ou pelo uso de algorítmo BLAST com o ácido nucleico ou sequência de polipeptídeo de interesse.

Exemplo 13: Alinhamento das sequências de polipeptídeos JMJC

O alinhamento das sequências de polipeptídeos foi realizado usando o programa AlignX a partir do vetor NTI (Invitrogen) que é baseado no algorítmo Clustal W popular do alinhamento progressivo (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31:3497-3500). Os valores default são pela penalidade aberta de fenda de 10, pela penalidade de extensão de fenda de 0,1 e a matriz do peso selecionado é Blosum 62 (se os polipeptídeos são alinhados). Uma edição menor do manual pode ser feita ainda para otimizar o alinhamento. A conservação da sequência entre os polipeptídeos JMJC é a maior no domínio JmjC dos polipeptídeos. A região correspondente aos motivos representados pela SEQ ID N°: 79 e pela SEQ ID N°: 81 é mais conservados do que do Motivo 8. Uma sequência de consenso é dada. Os resíduos de aminoácidos nas sequências de consenso são altamente conservados. Os brancos nas sequências conservadas representam qualquer aminoácido. Os polipeptídeos JMJC são alinhados na Figura 7.

Exemplo 14: Cálculo da identidade global da porcentagem entre as sequências de polipeptídeos JMJC úteis na realização de métodos da invenção

As porcentagens globais de similaridade e identidade entre as

sequências de polipeptídeos de comprimento total úteis na realização de métodos da invenção foram determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Ferramenta de alinhamento global de matriz) (BMC Bioinformatics. 2003 4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando sequências de DNA ou proteínas. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). O software MatGAT gera as matrizes de similaridade/identidade para sequência de proteínas ou DNA sem necessariamento o pré-alinhamento dos dados. O programa realiza uma série de alinhamento em pares usando o algorítmo de alinhamento global Myers and Miller (com uma penalidade de abertura de fenda de 12, e a penalidade de extensão de fenda de 2), calcula a similaridade e identidade usando por exemplo Blosum 62 (para os polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade da sequência é mostrada na metada da parte inferior da linha de divisão e a identidade de sequência é mostrada na metada da parte superior da linha de divisão diagonal.

Parâmetros usados na comparação foram:

Matriz de contagem:	Blosum62 219
Primeira fenda:	12
Extensão da fenda:	2

Os resultados da análise do software são mostrados na Tabela B2 pela similiaridade e identidade global no comprimento total de uma sequência de polipeptídeos. A identidade da porcentagem é dada acima em diagonal em negrito e a porcetagem da similaridade é dada abaixo em diagonal (face normal).

A identidade da porcentagem entre as sequências de polipeptídeos JMJC úteis na realização de métodos da invenção podem ser tão baixo quanto 15 % da identidade de aminoácido comparado a SEQ ID N°: 74.

Tabela B2: Resultados MatGAT para a similaridade e identidade global no

comprimento total de uma sequência de polipeptídeos.

		SEQ ID N0:74	SEQ ID N0:84	SEQ ID N0:86	SEQ ID N0:90	SEQ ID N0:92	SEQ ID N0:94	SEQ ID N0:136
AT3G20810.1	SEQ ID NO: 74		97,4	16,1	15,5	16,9	17,3	15,5
AT3G20810.2	SEQ ID NO: 84	97,4		15,5	14,5	17,5	17,1	14,9
AT5G19840	SEQ ID NO: 86	30,3	30,7		16,9	17	16,9	16,9
ABE92082	SEQ ID NO: 90	28,2	27,7	28,5		51,8	51,7	17,3
CAJ26373	SEQ ID NO: 92	30,1	29,6	27,7	67,6		82,5	12,4
Os09g0483600	SEQ ID NO: 94	30,1	30,1	27,5	67,3	89,6		16,1
Os01g36630	SEQ ID NO: 136	33,7	31,2	32,9	34,2	29,1	29,6	

Exemplo 15: A identificação dos domínios compreendidos nas sequências de polipeptídeos JMJC úteis na realização de métodos da invenção

O recurso integrado das famílias de proteínas, os bancos de dados de locais e domínios (InterPro) é uma interface integrada para os bancos de dados assinados comumente usados para as buscas com base no texto e sequência. O banco de dado InterPro combina estes bancos de dados, que usam as metodologias diferentes e variam os graus de informação biológica sobre as proteínas bem caracterizadas para derivar os sinais da proteína. A colaboração dos bancos de dados incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. O Pfam é uma ampla coleção de alinhamentos de sequência múltipla e modelos Markov secretos que cobrem muitos dos domínios de proteína comuns e famílias. O Pfam é hospedado no servidor Sanger Institute in the United Kingdom. O Interpro é hospedado no European Bioinformatics Institute in the United Kingdom.

A Tabela B3 mostra os ajustamentos (corte Gathering, corte de confiança e corte de ruído) como descrito nos bancos de dados Pfam que foram usados para produzir o HMMs_fs pelos diferentes domínios.

Tabela B3: Ajustamentos de HMMs_fs.

Domínio	Corte Gathering	Corte de confiança	Corte de ruído
PF02373	16	16,1	15,9
PF02375	15	16,8	13,8
PF02928	25	44,4	21,1
PF00646	17,7	17,7	17,6
PF00096	22,5	22,5	22,4
PF04967	15,4	15,4	3

Os resultados da avaliação Pfam para os polipeptídeos representativos JMJC de origem vegetal são apresentados na Tabela B4. As coordenadas do aminoácido para cada domínio na sequência de referência é indicado nas colunas de ínicio e fim. O valor E do alinhamento também é dado. O número de acessão InterPro ID de cada domínio identificado também é fornecido.

Tabela B4: Resultados de avaliação Pfam (maiores números de acessão) de polipeptídeos representativos JMJC de origem vegetal.

SEQ ID N° da proteína	Número (nome) de acessão Genbank (local)	banco de dados	entrada	Valor_E	início	fim	InterPro ID	descrição InterPro
96	AT1G08620	Pfam	PF02373	1.10E-62	368	484	1PR003347	Fator de transcrição jumonji JmjC
96	AT1 G08620	Pfam	PF02375	1.50E-16	127	165	IPR003349	Fator de transcrição jumonji JmjN
96	AT1 G08620	Pfam	PF02928	9.10E-22	591	644	IPR004198	Dedo de Zn, tipo C5HC2
96	AT1G08620	SMART	SM00541	1.10E-17	962	1006	1PR003888	FY-rich domain, N-terminal
96	AT1G08620	SMART	SM00542	1.30E-43	1012	1106	IPR003889	Domínio rico em FY, terminal N
98	AT1 G09060	Pfam	PF02373	6.80E-06	644	856	IPR003347	Fator de transcrição jumonji JmjC
104	AT1G11950	Pfam	PF02373	1.00E-23	645	825	1PR003347	Fator de transcrição jumonji jmjC
106	AT1G30810	Pfam	PF02373	5.00E-30	294	378	1PR003347	Fator de transcrição jumonji jmjC
106	AT1G30810	Pfam	PF02375	3.10E-24	58	104	1PR003349	Fator de transcrição jumonji jmjN
106	AT1G30810	Pfam	PF02928	4.50E-30	487	540	1PR004198	Dedo de Zn, tipo C5HC2
106	AT1G30810	SMART	SM00541	1.30E-17	626	670	1PR003888	Domínio rico em FY, terminal N
106	AT1G30810	SMART	SM00542	1.20E-39	676	762	1PR003889	Domínio rico em FY, terminal N
108	AT1G62310	Pfam	PF02373	2.00E-27	733	846	IPR003347	Fator de transcrição jumonji jmjC
108	AT1G62310	ProSite	PS50089	10.173	209	256	1PR001841	Zn-finger, RING

110	AT1G63490	Pfam	PF00628	5.40E-08	1009	1053	1PR001965	Semelhante a dedo de zinco, dedo de PHD
110	AT1G63490	Pfam	PF02373	3.10E-60	66	182	1PR003347	Fator de transcrição jumonji jmjC
110	AT1 G63490	Pfam	PF02928	4.80E-33	276	329	IPR004198	Dedo de Zn, tipo C5HC2
112	AT1G78280	Pfam	PF00646	0.00012	15	62	IPR001810	Cyclin-like F-box
112	AT1G78280	Pfam	PF02373	4.10E-12	249	362	IPR003347	Fator de transcrição jumonji jmjC
112	AT1G78280	ProSite	PS50181	9.603	14	60	IPR001810	F box semelhante à ciclina
114	AT2G34880	Pfam	PF02373	1.50E-67	294	410	1PR003347	Fator de transcrição jumonji jmjC
114	AT2G34880	Pfam	PF02375	3.70E-29	60	106	1PR003349	Fator de transcrição jumonji jmjN
114	AT2G34880	Pfam	PF02928	1.40E-23	514	567	1PR004198	Dedo de Zn, tipo C5HC2
114	AT2G34880	SMART	SM00541	9.90E-17	643	687	1PR003888	Domínio rico em FY, terminal N
114	AT2G34880	SMART	SM00542	4.60E-39	693	781	1PR003889	Domínio rico em FY, terminal N
116	AT2G38950	Pfam	PF02373	4.60E-47	321	437	1PR003347	Fator de transcrição jumonji jmjC
116	AT2G38950	Pfam	PF02375	9.30E-28	107	153	1PR003349	Fator de transcrição jumonji jmjN
116	AT2G38950	Pfam	PF02928	5.90E-27	544	597	1PR004198	Dedo de Zn, tipo C5HC2
118	AT3G07610	Pfam	PF02373	3.40E-15	726	843	1PR003347	Fator de transcrição jumonji jmjC
120	AT3G48430	Pfam	PF00096	0.014	1243	1268	1PR007087	Dedo de Zn, tipo C2H2
120	AT3G48430	Pfam	PF00096	0.0022	1296	1320	1PR007087	Dedo de Zn, tipo C2H2
120	AT3G48430	Pfam	PF00096	0.00026	1326	1352	1PR007087	Dedo de Zn, tipo C2H2
120	AT3G48430	Pfam	PF02373	9.00E-52	233	352	1PR003347	Fator de transcrição jumonji jmjC
120	AT3G48430	Pfam	PF02375	1.30E-07	19	62	1PR003349	Fator de transcrição jumonji jmjN
120	AT3G48430	ProSite	PS00028	0.00008	1268	1290	1PR007087	Dedo de Zn, tipo C2H2
122	AT4G00990	Pfam	PF02373	2.80E-16	607	781	IPR003347	jmjC Transcription factor jumonji,
124	AT4G20400	Pfam	PF02373	4.60E-68	239	355	1PR003347	Fator de transcrição

								jumonji jmjC
124	AT4G20400	Pfam	PF02375	9.60E-11	7	44	1PR003349	Fator de transcrição jumonji jmjN
124	AT4G20400	Pfam	PF02928	5.70E-30	462	515	1PR004198	Dedo de Zn, tipo C5HC2
124	AT4G20400	SMART	SM00541	4.40E-15	683	727	IPR003888	Domínio rico em FY, terminal N
124	AT4G20400	SMART	SM00542	7.40E-41	733	830	1PR003889	Domínio rico em FY, terminal N
128	AT5G04240	Pfam	PF00096	0.41	1228	1253	1PR007087	Dedo de Zn, tipo C2H2
128	AT5G04240	Pfam	PF00096	0.0028	1281	1305	1PR007087	Dedo de Zn, tipo C2H2
128	AT5G04240	Pfam	PF00096	0.0016	1311	1337	1PR007087	Dedo de Zn, tipo C2H2
128	AT5G04240	Pfam	PF02373	1.10E-48	292	411	1PR003347	Fator de transcrição jumonji jmjC
128	AT5G04240	Pfam	PF02375	1.00E-09	15	58	1PR003349	Fator de transcrição jumonji jmjN
130	AT5G06550	Pfam	PF00646	0.0013	81	128	1PR001810	F box semelhante à ciclina
130	AT5G06550	Pfam	PF02373	3.20E-07	311	422	IPR003347	Fator de transcrição jumonji jmjC
130	AT5G46910	Pfam	PF02373	1.00E-49	199	322	1PR003347	Fator de transcrição jumonji jmjC
132	AT5G46910	Pfam	PF02375	2.90E-09	21	62	IPR003349	Fator de transcrição jumonji jmjN
134	AT5G63080	Pfam	PF02373	3.90E-16	164	270	IPR003347	Fator de transcrição jumonji jmjC
74	AT3G20810	Pfam	PF02373	0.00013	378	409	IPR003347	Fator de transcrição jumonji jmjC
84	AT5G19840	Pfam	PF02373	0.036	173	281	1PR003347	Fator de transcrição jumonji jmjC
86	AT5G19840	Pfam	PF04967	0.47	295	307	1PR007050	HTH DNA binding domain
170	Gm_JMJ_1	Pfam	PF02373.11	9.70E-06	297	404	IPR003347	domínio JmjC

Exemplo 16: Clonagem da sequência JMJC de ácido nucleico usado nos métodos da invenção

A sequência do ácido nucleico usado nos métodos da invenção foi amplificado por PCR usando como modelo uma biblioteca de cDNA de mudas feitas comumente de *Arabidopsis thaliana* (em pCMV Sport 6.0;

Invitrogen, Paisley, UK). O PCR foi realizado usando a polimerase Hifi Taq DNA em condições padrões, usando 200 ng do modelo em uma mistura de PCR a 50 µl. Um iniciador a montante e a jusante como representado pela SEQ ID N°: 75 e SEQ ID N°: 76 respectivamente foram usados para amplificar por PCR (Reação de cadeia de polimerase) a região codificadora de JMJC como representado pela SEQ ID N°: 73. Os iniciadores incluem os locais AttB para a recombinação Gateway para facilitar a clonagem do fragmento de DNA de PCR amplificado no vetor de clonagem Gateway.

O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante o qual fragmento PCR recombina in vivo com o plasmídio pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada", pJMJC. O plasmídio pDONR201 foi adquirido de Invitrogen, como partes da tecnologia Gateway®.

O clone de entrada que compreende a SEQ ID N°: 73 foi então usado em uma reação LR com um vetor de destino usado pela transformação *Oryza sativa*. Este vetor contido como os elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; um cassete de expressão marcador avaliável; e um cassete Gateway pretendido pela recombinação LR in vivo com a sequência do ácido nucleico de interesse já clonado no clone de entrada. Um promotor GOS2 de arroz (SEQ ID N°: 77) para a expressão constitutiva foi localizada a montante deste cassete Gateway.

Após a etapa de recombinação LR, o vetor de expressão resultante pGOS2::JMJC (Figura 9) foi trasnformado na cepa da *Agrobacterium* LBA4044 de acordo com os métodos bem conhecidos na técnica.

Exemplo 17: Transformação da planta

A transformação das plantas foi realizada de acordo com os

procedimentos resumidos no Exemplo 7

Exemplo 18: Procedimento de avaliação fenotípica

18.1 Apresentação da avaliação

Aproximadamente 35 transformantes de arroz TO independentes foram gerados. Os transformantes primários foram transferidos a partir de uma câmera de cultura de tecido de uma estufa para o desenvolvimento e coletado da semente T1. Oito eventos, de que uma progênie T1 segregada 3:1 para a presença/ausência do transgene, foi retida. Para cada um destes eventos, aproximadamente 10 mudas T1 contendo o transgene (hetero e homozigotos) e aproximadamente 10 mudas T1 perderam o transgene (nulizigotos) foram selecionadas pela expressão marcadora visual do monitoramento. As plantas transgênicas e os nulizigotos correspondentes foram desenvolvidos lado a lado em posições aleatórias. As condições da estufa foram de dias curtos (12 horas de luz), 28° C na luz e 22° C no escuro, e uma umidade relativa de 70 %. O desenvolvimento das plantas sob as condições de não tensão são fornecidos com água em intervalos regulares para garantir que a água e nutrientes não são limitantes para satisfazer a necessidade da planta para completar o crescimento e desenvolvimento.

Quatro eventos T1 ainda foram avaliados na geração T2 seguindo o mesmo procedimento de avaliação como para a geração T1 mas com mais indivíduos por evento. A partir do estágio de semeadura até o estágio de maturidade das plantas foram passados diversas vezes através de um gabinete de imagem digital. Em cada período de tempo as imagens digitais (2048x1536 píxeis, 16 milhões de cores) foram tiradas de cada planta em pelo menos 6 ângulos diferentes.

Avaliação da aridez

As plantas a partir das sementes T2 foram desenvolvidas no solo em potes sob condições normais até estes aproximarem do estágio superior. Estes foram então transferidos à seção "seca" onde a irrigação foi

retida. As sondas de umidade foram inseridas em potes escolhidos aleatoriamente para monitorar o teor de água no solo (SWC). Quando o SWC está abaixo de certos limiares, as plantas foram automaticamente novamente submetidas à água continuamente até um nível normal foi atingido novamente. As plantas foram então re-transferidas novamente a condições normais. O resto do cultivo (amadurecimento da planta, coleta da semente) foi o mesmo para as plantas não desenvolvidas sob as condições de tensão abiótica. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais.

Avaliação da eficiência do uso de nitrogênio

As plantas de arroz das sementes T2 são desenvolvidas no solo em potes sob condições normais exceto para a solução do nutriente. Os potes são submetidos à água a partir da transplantação do amadurecimento com uma solução de nutriente específico contendo o teor de nitrogênio N reduzido (N), usualmente entre 7 a 8 vezes menor. O resto do cultivo (amadurecimento da planta, coleta da semente) é o mesmo como para as plantas não desenvolvidas sob a tensão abiótica. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais.

18.2 Análise estatística: Teste F

Dois fatores ANOVA (análise de variantes) foi usado como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformados com o gene da presente invenção. O teste F foi realizado para checar por um efeito do gene em todos os eventos de transformação e para verificar por um total do efeito do gene, também conhecido como um efeito do gene global. O limiar para a significância por um efeito verdadeiro do gene global foi apresentado em uma 5 % do nível de probabilidade para o teste F. Um ponto de valor do teste F significante por um

efeito do gene, significando que este não é apenas uma mera presença ou posição do gene que está causando as diferenças no fenotipo.

Por causa de dois experimentos com os eventos de sobreposição foram realizados, uma análise combinada foi realizada. Este é útil para checar a consistência dos efeitos nos dois experimentos, e se este é o caso, para acumular a evidência de ambos experimentos a fim de aumentar a confidência da conclusão. O método usado foi um método modelo misturado que leva em conta a estrutura multinível dos dados (isto é experimento - evento - segregantes). Os valores P obtidos em comparação ao teste de razão de probabilidade para distribuições quadradas *chi*.

18.3 Parâmetros medidos

Medição do parâmetro relacionado à biomassa

A partir do estágio de semeadura até o estágio de maturidade das plantas foram passados diversas vezes através de um gabinete de imagem digital. Em cada período das imagens digitais (2048x1536 píxeis, 16 milhões de cores) foram tiradas de cada planta em pelo menos 6 ângulos diferentes.

Uma área da planta acima do solo (ou biomassa da folha) foi determinado pela contagem do número total de píxeis nas imagens digitais a partir das partes de plantas acima do solo discriminadas a partir do fundamento. Este valor foi medido pelos desenhos retirados do mesmo ponto de tempo de ângulos diferentes e foi convertido a um valor de superfície física expressada em quadrado mm pela calibração. Os experimentos mostram que à área da planta acima do solo medida desta maneira correlaciona com a biomassa das partes das plantas trituradas acima. À área triturada acima é a área medida no ponto de tempo em que a planta tem atingido esta biomassa máxima da folha. O vigor precoce é à área das planta (mudas) acima do solo três semanas pós-germinação. O aumento na biomassa da raiz é expressado como um aumento no total da biomassa da raiz (medido como biomassa máxima de raízes observadas durante a vida toda de uma planta); ou como um

aumento na raiz/índice de broto (medido como a razão entre massa de raiz e massa de broto no período de desenvolvimento ativo da raiz e broto).

O vigor precoce foi determinado pela contagem do número total de píxeis a partir das partes de plantas acima do solo discriminadas a partir do fundamento. Este valor foi medido pelas figuras tiradas no mesmo tempo de ângulos diferentes e foi convertido a um valor de superfície física expressada em quadrado mm pela calibração. Os resultados descritos abaixo são para as plantas com três semanas pós-germinação.

Medições dos parâmetros relacionados à semente

As panículas primárias maduras foram coletadas, contadas, ensacadas, rotuladas com código de barra e então secadas por três dias em um forno a 37° C. As panículas foram então limiadas e todas as sementes foram coletadas e contadas. As cascas enchidas foram separadas a partir de um vazio usando um dispositivo de sopro de ar. As cascas vazias foram descartadas e a fração remanescente foi contada novamente. As cascas enchidas foram pesadas em uma balança analítica. O número de sementes enchidas foi determinado pela contagem do número de cascas enchidas que permanece depois da etapa de separação. A produção total de semente foi medida pela pesagem total das cascas enchidas coletadas a partir da planta. O número total de semente por planta foi medida pela contagem do número de cascas coletadas a partir da planta. Peso de núcleo milhar (TKW) é extrapolado do número de sementes enchidas contadas e seu peso total. O índice de colheita (HI) na presente invenção é definida como a razão entre a produção total de semente e à área triturada acima (mm^2), multiplicada por um fator 10^6. O número total de flores por panícula como definido na presente invenção é a razão entre o número total de sementes e o número de panículas primárias maduras. A taxa de abastecimento da semente como definido na presente invenção é a proporção (expressado como uma %) do número de sementes enchidas no número total de sementes (ou florzinhas).

Exemplo 19: Resultados da avaliação fenotípica das plantas transgênicas

Os resultados da avaliação das plantas transgênicas de arroz que expressam o ácido nucleico JMJC como representado pela SEQ ID N°: 73 sob as condições de não tensão são apresentadas abaixo. Um aumento em pelo menos 5 % foi observado para o vigor de emergência (vigor precoce), índice de broto/raiz, total da produção de semente, índice de colheita, e em pelo menos 3 % por peso de núcleo milhar nas plantas transgênicas quando comparados ao controle de plantas nulizigotos.

Os resultados da avaliação das plantas transgênicas de arroz que expressam a SEQ ID N°: 73 sob condições de tensão-aridez são apresentadas abaixo. Um aumento em pelo menos 5 % foi observado para o índice de broto/raiz, total peso da semente, número de sementes enchidas, taxa de abastecimento, índice de colheita e em pelo menos 3 % por peso de núcleo milhar nas plantas transgênicas quando comparados ao controle de plantas nulizigotos.

III. Caseína de quinase I

Exemplo 20: A identificação das sequências relacionadas à sequência CKI de ácido nucleico usado nos métodos da invenção

As sequências (cDNA eSTs de comprimento total ou genômico) relacionadas a uma sequência do ácido nucleico usadas nos métodos da presente invenção são identificados entre estes mantidos nos bancos de dados Entrez Nucleotides at the National Center for Biotechnology Information (NCBI) usando as ferramentas de sequência de banco de dados, tal como a ferramenta de alinhamento local básica (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa é usado para observar as regiões da similaridade local entre as sequências em comparação ao ácido nucleico ou sequências de polipeptídeos de um banco de dados da sequência e pelo cálculo da significância estatística de pontos. Por exemplo, o polipeptídeo codificado

pelos ácidos nucleicos usado na presente invenção são usados pelo algorítmo TBLASTN, com ajustes default e o filtro para ignorar baixa complexidade da sequência apresentada. A produção da análise é revisada pela comparação em pares, e classificada de acordo com as contagens de probabilidade (valor E), onde a contagem reflete a probabilidade que um alinhamento particular ocorra pela possibilidade (o valor E inferior, quanto mais significante o acerto). Além disso os valores E, em comparação são também contados pela identidade de porcentagem. A identidade da porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácido nucleico comparado (ou polipeptídeo) em um comprimento particular. Em alguns exemplos, os parâmetros default podem ser ajustados para modificar a estringência da busca. Por exemplo o valor E pode ser aumentado para mostrar menores pontos estringentes. Desta maneira, os pontos curtos próximos exatos podem ser identificados.

Em alguns exemplos, as sequências relacionadas podem ser tentativamente reunidas e publicadas, descobertas pelas instituições de busca, tal como The Institute for Genomic Research (TIGR). Os banco de dados Eukaryotic Gene Orthologs (EGO) podem ser usados para identificar tais sequência relacionadas pela busca de palavra chave ou pelo uso de algorítmo BLAST com o ácido nucleico ou sequência de polipeptídeo de interesse.

Exemplo 21: Alinhamento das sequências GRP de polipeptídeos

O alinhamento das sequências de polipeptídeos é realizado usando o programa AlignX a partir da embalagem do vetor NTI (Invitrogen) que é baseado no algorítmo Clustal W popular do alinhamento progressivo (Thompson et at. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31:3497-3500). Os valores default são pela penalidade aberta de fenda de 10, pela penalidade de extensão de fenda de 0,1 e a matriz do peso selecionado é Blosum 62 (se os polipeptídeos são alinhados). Uma edição menor do manual pode ser feita ainda para otimizar o

alinhamento.

A uma árvore filogenética de polipeptídeos GRP é construído usando um algorítmo de ajustamento de ligação próxima como fornecido no programa AlignX do vetor NTI (Invitrogen).

Exemplo 22: Cálculo da identidade global da porcentagem entre as sequências de polipeptídeos úteis na realização de métodos da invenção

As porcentagens globais de similaridade e identidade entre as sequências de polipeptídeos de comprimento total úteis na realização de métodos da invenção são determinadas usando um dos métodos disponíveis na técnica, o software MatGAT (Ferramenta de alinhamento global de matriz) (BMC Bioinformatics. 2003 4:29. MatGAT: uma aplicação que gera as matrizes de similaridade/identidade usando sequências de DNA ou proteínas. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). O software MatGAT gera as matrizes de similaridade/identidade para as sequência de proteínas ou DNA sem necessariamento o pré-alinhamento dos dados. O programa realiza uma série de alinhamento em pares usando o algorítmo de alinhamento global Myers and Miller (com uma penalidade de abertura de fenda de 12, e a penalidade de extensão de fenda de 2), calcula a similaridade e identidade usando por exemplo Blosum 62 (para os polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade da sequência é mostrada na metada da parte inferior da linha de divisão e a identidade de sequência é mostrada na metada da parte superior da linha de divisão diagonal.

Parâmetros usados na comparação foram:

Matriz de contagem: Blosum62

Primeira fenda: 12

Extensão da fenda: 2

A Tabela MATGAT para o alinhamento local de um domínio específico, ou dados em % da identidade/similarity entre os domínios

específicos também podem ser gerados.

Exemplo 23: A identificação dos domínios compreendidos nas sequências de polipeptídeos úteis na realização de métodos da invenção

O recurso integrado das famílias de proteínas, bancos de dados de locais e domínios (InterPro) é uma interface integrada para os bancos de dados assinados comumente usados para as buscas com base no texto e sequência. O banco de dado InterPro combina este banco de dados, que usam as metodologias diferentes e variam os graus de informação biológica sobre as proteínas bem caracterizadas para derivar os sinais da proteína. A colaboração dos bancos de dados incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. O Pfam é uma ampla coleção de alinhamentos de sequência múltipla e modelos Markov secretos que cobrem muitos dos domínios de proteína comuns e famílias. O Pfam é hospedado no servidor do Sanger Institute in the United Kingdom. O Interpro é hospedado no European Bioinformatics Institute in the United Kingdom.

A sequência de proteínas que representam o GRP são usadas como questão para buscar o banco de dado InterPro.

Exemplo 24: Predição da topologia de uma sequência de polipeptídeos úteis na realização de métodos da invenção

O TargetP 1.1 prediz a localização subcelular das proteínas eucarióticas. A indicação da localização é com base na presença predita de qualquer uma das pré-sequências do terminal N: peptídeo de transição de cloroplasto (cTP), peptídeo de alvejamento mitocondrial (mTP) ou peptídeo de sinal de caminho secretório (SP). As contagens em que a predição final não são baseados nas reais probabilidades, e estes não são necessariamente adicionado a um. Entretanto, a localização com a contagem mais alta é igualmente de acordo com o TargetP, e a conexão entre as contagens (a classe de confiabilidade) pode ser uma indicação de que certas predições são. A

classe de confiabilidade (RC) varia de 1 a 5, onde 1 indica a predição mais forte. O TargetP é mantido no servidor da Technical University of Denmark.

Para as sequências preditas por conter uma pré-sequência de terminal N e um local de clivagem potencial também pode ser predito.

Um número de parâmetros foram selecionados, tal como grupo de organismos (vegetal ou não vegetal), séries de cortes (nenhum, série pré-definida de cortes, ou série especificadas pelo uso de cortes), e o cálculo da predição dos locais de clivagem (sim ou não).

A sequência de proteínas que representam o GRP são usadas para questionar o TargetP 1.1. O grupo de organismos "vegetal" é selecionado, nos cortes definidos, e o comprimento predito do peptídeo de transição requerido.

Muitos outros algorítmos podem ser usados para realizar tais análises, incluindo:

- ChloroP 1.1 hospedado no servidor da Technical University of Denmark;
- Protein Prowler Subcellular Localisation Predictor versão 1.2 hospedado no servidor do Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia;
- PENCE Proteome Analyst PA-GOSUB 2.5 hospedado no servidor da University of Alberta edmonton, Alberta, Canada;
- TMHMM, hospedado no servidor da Technical University of Denmark

Exemplo 25: Clonagem da sequência do ácido nucleico usado nos métodos da invenção

Clonagem da SEQ ID N°: 171:

A sequência do ácido nucleico da SEQ ID N°: 171 usada nos métodos da invenção foi amplificado por PCR usando como modelo uma feitas comumente das biblioteca de cDNA de mudas *Arabidopsis thaliana*

(em pCMV Sport 6.0; Invitrogen, Paisley, UK). O PCR foi realizado usando a polimerase Hifi Taq DNA em condições padrões, usando 200 ng do modelo em uma mistura de PCR a 50 µl. Os iniciadores usados foram prm8667 (SEQ ID N°: 177; sentido, códon de partida em negrito):

5'-ggggacaagtttgtacaaaaaagcaggcttaaacaatggactgcaacatggtatct-3'

e prm8668 (SEQ ID N°: 178; reverso, complementar):

5'-ggggaccactttgtacaagaaagctgggtcacattacttactcatctattttgg-3',

que incluem os locais AttB para a recombinação Gateway. O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante o qual fragmento PCR recombina in vivo com o plasmídio pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada". O plasmídio pDONR201 foi adquirido de Invitrogen, como partes da tecnologia Gateway®.

Clonagem da SEQ ID N°: 173:

Um experimento cDNA-AFLP foi realizado em uma cultura celular BY2 de tabaco sincronizado *(Nicotiana tabacum* L. *cv.* Bright Yellow-2), e rótulos de sequência expressados BY2 que foram modulados pelo ciclo celular e foram eleitos para a clonagem adicional. Os rótulos de sequência expressados foram usados para avaliar uma biblioteca de cDNA de tabaco e para isolar o cDNA de comprimento total de interesse, nomeado um codificador para a SEQ ID N°: 173.

Uma suspensão celular cultivada de tabaco BY2 *(Nicotiana tabacum* L. *cv.* Bright Yellow-2) foi sincronizada pelas células que bloqueiam na fase S precocimente como afidicolina como seguem. A suspensão celular de *Nicotiana tabacum* L. cv. Bright Yellow 2 foi mantida como descrito (Nagata et al. Int. Rev. Cytol. 132, 1-30, 1992). Para a sincronização, uma cultura estacionária de 7 dias de idade foi diluída 10 vezes em meio fresco

suplementado com afidicolina (Sigma-Aldrich, St. Louis, MO; 5 mg/l). Após 24 horas, as células foram liberadas do bloqueio por diversas lavagens com meio fresco após ter o seu ciclo celular de progressão resumido.

O RNA total foi preparado usando a precipitação LiCI e RNA poli(A+) foi extraído de 500 μg do RNA total usando colunas Oligotex (Qiagen, Hilden, Germany) de acordo com as instruções do fabricante. Partida de 1 μg de RNA poli(A+), primeiro cDNA de filamento foi sintetizado pela transcrição reversa com um iniciador oligo-dT_{25} biotinilado (Genset, Paris, France) e Superscript II (Life Technologies, Gaithersburg, MD). A segunda síntese de filamento foi feita pelo deslocamento do filamento com ligase *Escherichia* coli (Life Technologies), polimerase de DNA I (USB, Cleveland, OH) e RNAse-H (USB).

Quinhentos ng de cDNA filamentado duplo foi usado pela análise AFLP como descrito (Vos et al., Nucleic Acids Res. 23 (21) 4407-4414, 1995; Bachem et al., Plant J. 9 (5) 745-53, 1996) com modificações. As enzimas de restrição usadas foram BstYl e Msel (Biolabs) e a digestão foi feita em duas etapas separadas. Após a primeira restrição da digestão com uma das enzimas, os fragmentos da extremidade final 3' foram presos naas pérolas Dyna (Dynal, Oslo, Norway) por meio de sua calda biotinilada, enquanto os outros fragmentos foram logo lavados. Após a digestão com a segunda enzima, os fragmentos de restrição liberados foram coletadas e usados como modelos nas etapas AFLP subsequentes. Para as pré-amplificações, um iniciador Msel sem nucleotídeo seletivos foi combinado com um iniciador BstYl contendo um T ou um C como mais nucleotídeos 3'. As condições PCR foram como descritos (Vos et al., 1995). As misturas de amplificação obtida foram diluídas 600 vezes e 5 μl foi usado para as amplificações seletivas usando um iniciador BstYl rotulado por P^{33} a polimerase Amplitaq-Gold (Roche Diagnostics, Brussels, Belgium). Os produtos de amplificação foram separados em 5 % de géis poliacrilamidas

usando o sistema Sequigel (Biorad). Géis secos foram expostos às películas Kodak Biomax bem como escaneadas em um PhosphorImager (Amersham Pharmacia Biotech, Little Chalfont, UK).

Os bandos correspondentes aos transcriptos diferentemente expressados, entre o qual o transcripto correspondente (parcial) da SEQ ID N°: 173, foram isolados a partir de gel e DNA eluído foi amplificado novamente sob as mesmas condições como para a amplificação seletiva. A informação da sequência foi obtida pelo sequenciamento direto da reação amplificada novamente de cadeia de produto de polimerase com o iniciador BstYI seletivo ou após a clonagem do fragmento em pGEM-T leve (Promega, Madison, WI) e sequenciamento de clones individuais. As sequências obtidas foram comparadas contra os nucleotídeos e as sequências de proteínas presentes nos bancos de dados publicamente disponíveis pelos alinhamento de sequência BLAST (Altschul *et al.,* Nucleic Acids Res. 25 (17) 3389-3402 1997). Quando disponível, as sequências de rótulo foram substituídas com EST mais longos ou sequências de cDNA isolados para aumentar a chance de observar a homologia significante. O clone cDAN físico correspondente a SEQ ID N° 173 foi subsequentemente amplificado a partir de uma biblioteca de cDNA de tabaco comercial como seguem:

Uma biblioteca de c-DNA com um tamanho médio de inserções de 1.400 bp foi preparado de RNA poli(A+) isolado da divisão ativamente, nas células de tabaco BY2 não sincronizadas. Estas inserções na biblioteca foram clonadas no vetor pCMVSPORT6.0, que compreende um cassete Gateway attB (Life Technologies). A partir desta biblioteca, 46.000 clones foram selecionados, submetidos ao ensaio em plas microtituladoras de 384 reservatórios, e subsequentemente manchadas em reprodução em filtro de náilon. Os clones submentidos ao ensaio foram avaliados usando grupos de diversas centenas de rótulos rotulados radioativamente como sondas (incluindo o BY2-tag correspondente a uma sequência SEQ ID N°: 173). Os

clones positivos foram isolados (entre o qual o clone correspondente a SEQ ID N°: 173), sequenciado, e alinhado com a sequência do rótulo. Onde a hidridização com o rótulo falha, o cDNA de comprimento total correspondente ao rótulo foi selecionado pela amplificação de PCR: iniciadores específico de rótulo foram projetados usando software comumente disponível e usado em combinação com um iniciador de vetor comum para amplificar as inserções de cDNA parciais. Os grupos de DNA de 50.000, 100.000, 150.000, e 300.000 clones de cDNA foram usados como modelos nas amplificações de PCR. Os produtos de amplificação foram então isolados a partir dos géis de agarose, clonado, sequenciado e sua sequência alinhada com aqueles dos rótulos. Próximo, o cDNA de comprimento total correspondente à sequência de nucleotídeo da SEQ ID N° 173 foi clonada a partir do vetor da biblioteca pCMVsport6.0 em pDONR201, um vetor doador de Gateway® (Invitrogen, Paisley, UK) por intermédio de uma reação LR, resultando em um clone de entrada.

O clone de entrada que compreende a SEQ ID N°: 171 ou SEQ ID N°: 173 foi então usado em uma reação LR com um vetor de destino usado pela transformação de *Oryza sativa*. Este vetor contido como os elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; um cassete de expressão marcador avaliável; e um cassete Gateway pretendido pela recombinação LR in vivo com uma sequência do ácido nucleico de interesse já clonado no clone de entrada. Um promotor WSI18 de arroz (SEQ ID N°: 175) para a expessão específica da semente foi localizada a montante deste cassete Gateway.

Após a etapa de recombinação LR, o vetor de expressão resultante pWSI18::GRP (Figura 11) foi transformado na cepa de *Agrobacterium* LBA4044 de acordo com os métodos bem conhecidos na técnica.

Em uma construção alternativa do promotor GOS2 de arroz

(SEQ ID N°: 176) foi usado, resultando no vetor de expressão pGOS2::GRP.

Exemplo 26: Transformação da planta

A transformação das plantas foi realizada de acordo com os procedimentos resumidos no Exemplo 7.

Exemplo 27: Procedimento de avaliação fenotípica

27.1 Apresentação da avaliação

Aproximadamente 35 transformantes de arroz TO independentes foram gerados. Os primeiros transformantes foram transferidos a partir de uma câmera de cultura de tecido de uma estufa para o desenvolvimento e coletado da semente T1. Seis eventos, de que uma progênie T1 segregada 3:1 para a presença/ausência do transgene, foram retidas. Para cada um destes eventos, aproximadamente 10 mudas T1 contendo o transgene (hetero e homozigotos) e aproximadamente 10 mudas T1 perderam o transgene (nulizigotos) foram selecionadas pela expressão marcadora visual do monitoramento. As plantas transgênicas e os nulizigotos correspondentes foram desenvolvidos lado a lado em posições aleatórias. As condições da estufa foram de dias curtos (12 horas de luz), 28° C na luz e 22° C no escuro, e uma umidade relativa de 70 %. O desenvolvimento das plantas sob as condições de não tensão são fornecidos com água em intervalos regulares para garantir que a água e nutrientes não são limitantes para satisfazer a necessidade das plantas para completar o crescimento e desenvolvimento.

Quatro eventos T1 ainda foram avaliados na geração T2 seguindo o mesmo procedimento de avaliação como para a geração T1 mas com mais indivíduos por evento. A partir do estágio de semeadura até o estágio de maturidade das plantas foram passados diversas vezes através de um gabinete de imagem digital. Em cada período de tempo imagens digitais (2048x1536 píxeis, 16 milhões de cores) foram tiradas de cada planta em pelo menos 6 ângulos diferentes.

Avaliação da aridez

As plantas das sementes T2 foram colocadas em solo em potes sob condições normais até estas aproximarem do estágio superior. Estas foram então transferidas a uma seção "seca" onde a irrigação foi retida. As sondas de umidade foram inseridas em potes escolhidos aleatoriamente para monitorar o teor de água no solo (SWC). Quando SWC está abaixo de certos limiares, as plantas foram atutomaticamente novamentes submetidas à água continuamente até um nível normal ser atingido novamente. As plantas foram então re-transferidas novamente a condições normais. O resto do cultivo (amadurecimento da planta, coleta da semente) foi a mesma como para as plantas não desenvolvidas sob as condições de tensão abiótica. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais.

Avaliação da eficiência do uso de nitrogênio

As plantas de arroz das sementes T2 são desenvolvidas no solo em potes sob condições normais exceto para a solução do nutriente. Os potes são submetidos à água a partir da transplantação do amadurecimento com uma solução de nutriente específico contendo o teor de nitrogênio N reduzido (N), usualmente entre 7 a 8 vezes menor.

O resto do cultivo (amadurecimento da planta, coleta da semente) é o mesmo como para as plantas não desenvolvidas sob a tensão abiótica. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais.

27.2 Análise estatística: Teste F

Dois fatores ANOVA (análise de variantes) foram usados como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformados com o gene da presente invenção. O teste F foi realizado para checar por um efeito do gene

em todos os eventos de transformação e para verificar por um total do efeito do gene, também conhecido como um efeito do gene global. O limiar para a significância por um efeito verdadeiro do gene global foi apresentado em uma 5 % do nível de probabilidade para o teste F. Um ponto de valor do teste F significante por um efeito do gene, significando que este não é apenas uma mera presença ou posição do gene que está causando as diferenças no fenotipo.

Por causa de dois experimentos com os eventos de sobreposição foram realizados, uma análise combinada foi realizada. Este é útil para checar a consistência dos efeitos nos dois experimentos, e se este é o caso, para acumular a evidência de ambos experimentos a fim de aumentar a confidência da conclusão. O método usado foi um método modelo misturado que leva em conta à estrutura multinível dos dados (isto é experimento - evento - segregantes). Os valores P obtidos em comparação ao teste de razão de probabilidade para distribuições quadradas chi.

27.3 Parâmetros medidos

Medição do parâmetro relacionado à biomassa

A partir do estágio de semeadura até o estágio de maturidade das plantas foram passados diversas vezes através de um gabinete de imagem digital. Em cada período de tempo imagens digitais (2048x1536 píxeis, 16 milhões de cores) foram tiradas de cada planta em pelo menos 6 ângulos diferentes.

Uma área da planta acima do solo (ou biomassa da folha) foi determinado pela contagem do número total de píxeis nas imagens digitais a partir das partes de plantas acima do solo discriminadas a partir do fundamento. Este valor foi medido pelas figuras retiradas do mesmo ponto de tempo a partir de ângulos diferentes e foi convertido a um valor de superfície física expressada em quadrado mm pela calibração. Os experimentos mostram que à área da planta acima do solo medida desta maneira correlaciona com a

biomassa das partes das plantas trituradas acima. À área triturada acima é a área medida no ponto de tempo em que a planta tem atingido esta biomassa máxima da folha.

O vigor precoce é à área das planta (mudas) acima do solo três semanas pós-germinação. O vigor precoce foi determinado pela contagem do número total de píxeis a partir das partes de plantas acima do solo discriminadas a partir do fundamento. Este valor foi medido pelas figuras tiradas no mesmo ponto de tempo de ângulos diferentes e foi convertido a um valor de superfície física expressada em quadrado mm pela calibração. Os resultados descritos abaixo são para as plantas com três semanas pós-germinação.

Medições dos parâmetros relacionados à semente

As panículas primárias maduras foram coletadas, contadas, ensacadas, rotuladas com código de barra e então secadas por três dias em um forno a 37° C. As panículas foram então limiadas e todas as sementes foram coletadas e contadas. As cascas enchidas foram separadas a partir de um vazio usando um dispositivo de sopro de ar. As cascas vazias foram descartadas e a fração remanescente foi contada novamente. As cascas enchidas foram pesadas em uma balança analítica. O número de sementes enchidas foi determinado pela contagem do número de cascas enchidas que permanece depois da etapa de separação. A produção total de semente foi medida pela pesagem total das cascas enchidas coletadas a partir da planta. O número total de semente por planta foi medido pela contagem do número de cascas coletadas a partir da planta. Peso de núcleo milhar (TKW) é extrapolado do número de sementes enchidas contadas e seu peso total. O índice de colheita (HI) na presente invenção é definida como a razão entre a produção total de semente e à área triturada acima (mm^2), multiplicada por um fator 10^6. O número total de flores por panícula como definido na presente invenção é a razão entre o número total de sementes e o número de panículas primárias

maduras. A taxa de abastecimento da semente como definido na presente invenção é a proporção (expressado como uma %) do número de sementes enchidas no número total de sementes (ou florzinhas).

Exemplo 28: Resultados da avaliação fenotípica das plantas transgênicas

As plantas transgênicas de arroz que expressam o ácido nucleico GRP representado pela SEQ ID N°: 171 sob o controle do promotor WSI18 mostrando um aumento de mais do que 5 % para a biomassa, peso total das sementes, número de sementes enchidas, taxa de abastecimento, e índice de colheita, quando desenvolvidos sob condições de tensão à aridez. Quando o desenvolvimento sob as condições de não tensão, existe um aumento de mais do que 5 % observado para a biomassa, número de sementes enchidas, peso total das sementes, e número das primeiras panículas.

Para a construção com a SEQ ID N°: 171 sob o controle do promotor GOS2, ujm aumento foi observado nas plantas transgênicas para o vigor precoce e para as flores por panícula, e para cada um destes parâmetros, o aumento foi maior do que 5 %.

As plantas transgênicas que expressam a SEQ ID N°: 173 sob o controle do promotor WSI18 e desenvolvidos sob as condições de tensão à aridez, mostrando um aumento no peso total das sementes, número de sementes enchidas, número de flores por panícula, índice de colheita, e peso de núcleo milhar.

Por peso de núcleo milhar o aumento observado foi em pelo menos 3,5 % e para os outros parâmetros o aumento foi maior do que 5 %.

IV. Polipeptídeo de prolongamento-homeodomínio de homeodomínio de planta (PHDf-HD)

Exemplo 29: A identificação das sequências relacionadas a uma sequência do ácido nucleico usado nos métodos da invenção

As sequências (cDNA eSTs de comprimento total ou

genômico) relacionadas a uma sequência do ácido nucleico usado nos métodos da presente invenção onde identificados entre aqueles mantidos no banco de dados Entrez Nucleotides at the National Center for Biotechnology Information (NCBI) usando as ferramentas de sequência de banco de dados, tal como a ferramenta de alinhamento local básica (BLAST) (Altschul et at. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa é usado para observar as regiões da similaridade local entre as sequências em comparação a sequência de ácido nucleico ou as sequências de polipeptídeos de um banco de dados da sequência e pelo cálculo da significância estatística de pontos. Por exemplo, o polipeptídeo codificado por uma sequência do ácido nucleico da presente invenção foi usado pelo algorítmo TBLASTN, com ajustes default e o filtro para ignorar baixa complexidade sequência apresentada. A produção da análise foi examinada pela comparação em pares, e classificada de acordo com as contagens de probabilidade (Valor e), onde a contagem reflete a probabilidade que um alinhamento particular ocorra pela possibilidade (o valor E inferior, quanto mais significante o acerto). Além disso aos valores E, as comparações também foram contadas pela identidade de porcentagem. A identidade da porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências comparadas de sequências de ácido nucleico (ou polipeptídeo) em um comprimento particular. Em alguns exemplos, os parâmetros default podem ser ajustados para modificar a estringência da busca. Por exemplo o valor E pode ser aumentado para mostrar os menores pontos estringentes. Desta maneira, os pontos curtos próximos exatos podem ser identificados.

A Tabela D1 fornece uma lista das sequências de ácidos nucleicos relacionados à sequência de ácido nucleico PHDf-HD usados nos métodos da presente invenção.

Tabela D1: Exemplos dos polipeptídeos PHDf-HD:

Nome	fonte de organismo	SEQ ID N° do ácido nucleico	SEQ ID N° de polipeptídeo	acessão do banco de dados #	condição
Orysa_PHDf_HD	*Oryza sativa*	179	180	Os02g05450.1 NM_001052422	FL
Arath_PHDf_HD_PRHA	*Arabidopsis thaliana*	181	182	At4g029940 NM_119140	FL
Eucgr_PHDf_HD	*Eucalyptus grandis*	183	184	ADW17964	FL
Medtr_PHDf_HD	*Medicago truncatula*	185	186	AC123547	FL
Pinra PHDf HD	*Pinus radiata*	187	188	ADW18458	FL
Poptr_PHDf_HD	*Populus tremuloides*	189	190	scaff_VI.625	FL
Sacof_PHDf_HD	*Saccharin officinarum*	191	192	CA157855.1, CA261734.1, CA253314.1, CA220753.1, CA201958.1	FL
Vitvi_PHDf HD	*Vitis vinifera*	193	184	AM477372.2, AM488059.1	FL
Zeama_PHDf HD	*Zea mays*	185	186	EE162310, DN204182, CF057937	FL
Arath_PHDf_HD_HAT3.1	*Arabidopsis thaliana*	187	188	AT3G19510 NM_112838	FL
Lotja_PHDf_HD	*Lotus japonicus*	189	190	AP006117.1	FL
Orysa_PHDf_HD_HAZ1	*Oryza sativa*	191	192	AB081340 0s06g 12400.1	FL
Petcr_PHDf HD_PRHP	*Petroselinum crispum*	193	194	L21975	FL
Zeama_PHDf_HD_HOX1*a*	*Zea mays*	195	196	X67561	FL
Zeama_PHDf_HD_HOX1b	*Zeamays*	197	198	X92428	FL
Zeama_PHDf_HD_HOX2a	*Zea mays*	199	200	X89760.1	FL
Zeama_PHDf HD_HOX2b	*Zea mays*	201	202	X89761	FL
Vitvi PHDf HD II	*Vitis vinifera*	203	204	AM464161.2 AM478203.2	FL
Poptr_PHDf_HD_II	*Populus tremuloides*	205	206	scaff_IX.730	FL
Poptr_PHDf_HD_III	*Populus tremuloides*	207	208	LG_1002624	FL
Poptr_PHDf_HD_IV	*Populus tremuloides*	209	210	LG_XVIII1192	FL
Ostta PHDf HD	*0streococcus tauri*	211	212	CR954214.4	FL
Aqufor_PHDf HD parcial	*Aquilegia formosa x Aquilegia pubescens*	213	214	DR914726.1, DR941696.1, DR943570.1	Parcial
Glyma_PHDf_HD	*Glycine max*	215	216	Contig GM06LC25006	Parcial
Lotco PHDf HD	*Lotus comiculatus*	217	218	AP004517	Parcial
Sorpr_PHDf_HD 3'	*Sorghum propinquum*	219	220	BF656332	Parcial
Sorpr_PHDf HD 5'	*Sorghum propinquum*	221	222	BF704605	Parcial
Phypa_PHDf_HD	*Physcomitrella*	238	239	XM_001762483	FL

	patens				
Phypa_PHDf HD	*Physcomitrella patens*	240	241	XM_001779822	FL
Zeama_PHDf_HD	*Zea mays*	242	243		FL

Exemplo 30: Alinhamento das sequências PHDf-HD de polipeptídeos

A University of Potsdam, Germany, tem criado banco de dados do fator de transcripção da planta, incluindo pela Orysa sativa nomeada TFDB de arroz. As sequências de polipeptídeos correspondentes aos fatores de transcripção pertencentes à família HD (120 modelos de genes (91 loci) identificado próximo) foram todos carregados, incluindo os dois polipeptídeos PHDf-HD (Os02g05450.1 e Os06g12400.1) identificado aos dados. As sequências de polipeptídeos da Tabela D1 da presente aplicação foram adicionados (quando o comprimento total, isto é 21 sequências de polipeptídeos) à série da família HD.

O alinhamento de todas as sequência de polipeptídeos foi realizado o algorítmo Clustal (1.83) do alinhamento progressivo, usando valores default (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31:3497-3500). A árvore de ligação próxima foi construída depois, e é representada na Figura 13 da presente aplicação. O grupo de interesse, que compreende os dois parálogos de arroz (Os02g05450.1 e Os06g12400.1) foram circulados. Qualquer polipeptídeo diminuído dentro deste grupo HD (após uma nova etapa de alinhamento múltiplo como descrito anteriomente acima) é considerado ser útil na realização de métodos da invenção como descrito neste.

Em um alinhamento de sequência múltipla dos polipeptídeos PHDf-HD de comprimento total da Tabela D1, um número de características podem ser identificadas, e são marcadas na Figura 17. A partir do terminal N ao terminal C dos polipeptídeos são: (i) um sinal de localização nuclear predito (NLS); (ii) um zíper de leucina (ZIP), com quatro *heptads* (em compartimentos em que usualmente uma leucina (ocasionalmente uma isoleucina, uma valina, ou uma metionina)) aparece a cada sétimo

aminoácido; (iii) um dedo PHD (PHDf), com o C4HC3 típico (quatro cisteínas, uma histidina, três cisteínas) com um espaçamento da cisteína característica; (iv) uma extensão ácida (rico em aminoácidos ácidos D e E); (v) extensões básicas (rica em aminoácidos básicos K e R); (vi) um homeodomínio (HD).

Exemplo 31: Cálculo da identidade global da porcentagem entre as sequências de polipeptídeos úteis na realização de métodos da invenção

As porcentagens globais de similaridade e identidade entre as sequências de polipeptídeos de comprimento total úteis na realização de métodos da invenção foram determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Ferramenta de alinhamento global de matriz) (BMC Bioinformatics.2003 4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando sequências de proteína ou DNA. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). O software MatGAT gera as matrizes de similaridade/identidade para sequência de proteínas ou DNA sem necessidade do pré-alinhamento dos dados. O programa realiza uma série de alinhamento em pares usando o algorítmo de alinhamento global Myers and Miller (com uma penalidade de abertura de fenda de 12, e a penalidade de extensão de fenda de 2), calcula a similaridade e identidade usando por exemplo Blosum 62 (para os polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade da sequência é mostrada na metada da parte inferior da linha de divisão e a identidade de sequência é mostrada na metada da parte superior da linha de divisão diagonal.

Parâmetros usados na comparação foram:

Matriz de contagem: Blosum62

Primeira fenda: 12

Extensão da fenda: 2

Os resultados da análise do software são mostrados na Tabela

D2 pela similiaridade e identidade global no comprimento total de uma sequência de polipeptídeos (excluindo as sequências parciais de polipeptídeos). A identidade da porcentagem é dada acima em diagonal e a porcetagem da similaridade é dada abaixo em diagonal.

A identidade da porcentagem entre as sequências PHDf-HD de polipeptídeo úteis na realização de métodos da invenção podem ser tão baixo quanto 15 % da identidade de aminoácido comparado a SEQ ID N°: 180.

Tabela D2: Resultados MatGAT para a similaridade e identidade global no comprimento total de uma sequência de polipeptídeos.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. Arath PHDf HD PRHA		25	31	40	19	38	26	17	32	40	20	19	40	30	40	23	31	25	25	18	19
2. Arath PHDf-HD HAT3	44		23	27	33	26	34	28	27	26	34	35	27	23	28	34	24	36	35	18	20
3. Orysa_PHDf_HD	46	39		35	18	34	20	18	31	34	17	18	33	56	36	21	57	21	23	15	15
4. Eucgr_PHDf_HD	57	45	51		21	47	27	20	36	50	22	21	51	32	52	25	32	26	27	17	17
5. Lotja_PHDf-HD	36	47	31	35		22	29	32	23	20	39	40	20	18	21	34	18	31	32	20	22
6. Medtr P H Df H D	55	44	51	64	36		26	19	35	50	20	20	49	30	50	25	32	27	26	17	17
7. Orysa_PHDf-HD_HAZ1	42	50	35	47	45	43		26	25	25	29	29	26	21	26	32	21	50	46	29	30
8. Petcr PHDf HD PRHP	31	39	28	32	48	31	38		22	20	32	33	20	18	20	30	18	26	25	19	21
9. Pinra P H D f H D	53	44	47	51	42	50	44	39		36	23	23	37	28	39	26	30	25	27	19	19
10. Poptr_PHDf_HD_I	58	45	52	67	36	66	45	33	52		20	21	84	30	55	25	31	25	25	17	17
11. Poptr_PHDf_HD_III	36	47	30	34	56	35	44	46	41	36		80	21	19	22	37	18	31	31	20	20
12. Poptr_PHDf_HD_II	36	49	31	34	57	34	46	47	42	35	86		21	18	22	38	18	30	30	20	21
13. Poptr_PHDf_HD_IV	59	46	51	68	37	67	44	33	54	91	35	35		30	57	25	32	25	25	17	17
14. Sacof PHDf HD	48	43	69	51	31	50	38	29	45	51	34	33	51		34	22	83	21	22	15	16
15. Vitvi PHDf HD I	58	47	52	69	36	67	43	32	52	74	36	35	75	52		23	34	27	25	16	17
16. Vitvi PHDf HD II	39	47	36	39	56	39	45	47	43	40	56	57	40	36	41		22	32	31	23	24
17. Zeama PHDf HD	48	41	70	51	32	52	38	30	45	54	33	32	53	87	53	34		22	22	15	15
18. Zeamahox1a	43	54	39	44	43	44	63	38	43	44	44	43	45	42	43	46	42		83	24	24
19. Zeamahox1b	42	53	40	44	44	42	60	36	43	44	44	43	45	41	44	45	41	88		23	23
20. Zeamahox2a	27	27	24	25	32	25	36	31	31	26	30	31	27	25	26	34	24	31	30		81
21. Zeama_Hox2b	29	28	23	27	33	27	37	34	31	27	32	33	26	25	27	35	25	32	30	86	

A identidade da porcentagem pode ser substancialmente aumentada se o cálculo da identidade é realizada entre o domínio ZIP/PHDf conservado (zíper de leucina conservado/domínio de dedo homeodomínio de planta, que compreende os domínios Zip e PHDf) da SEQ ID N°: 180 (como representado pela SEQ ID N°: 233) e o domínio ZIP/PHDf conservado dos polipeptídeos utéis na realização da invenção. O ZIP/PHDf conservado da SEQ ID N°: 233 é um total de 180 aminoácidos contígos de comprimento. A identidade da porcentagem no domínio ZIP/PHDf conservado entre a

sequência de polipeptídeos úteis na realização de métodos da invenção que variam entre 30 % e 75 % da identidade de aminoácido, como mostrado na Tabela D3.

Tabela D3: Resultados MatGAT para a similaridade e identidade global no domínio ZIP/PHDf conservado entre uma sequência de polipeptídeos.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	1*9	20	21
1. ZIP/PHDf_Arath_PHDf_HD_PRHA		52	**50**	77	48	74	53	48	64	73	48	46	72	51	76	50	52	50	50	53	53
2. ZIP/PHDf ArathPHDf-HD HAT3	70		**44**	52	73	54	70	69	56	52	75	75	51	**43**	53	78	46	65	67	64	65
3. ZIP/PHDf_Orysa_PHDf_HD	69	60		**56**	**39**	**54**	**37**	**39**	**56**	**53**	**39**	**38**	**53**	**70**	**59**	**41**	**71**	**37**	**38**	**34**	**33**
4. ZIP/PHDf_Eucgr_PHDf_HD	90	70	73		52	79	48	50	71	82	51	49	81	55	84	52	56	48	48	49	48
5. ZIP/Lotja_PHDf-HD	67	89	59	69		52	64	69	52	47	76	76	47	37	50	78	38	64	66	61	61
6. ZIP/Medtr_PHDf_HD	86	70	72	87	68		52	49	66	74	49	48	75	50	79	51	53	51	51	52	52
7. ZIP/PHDf_Orysa_PHDf-HD_HAZ1	69	83	54	67	83	68		64	51	47	68	67	47	35	47	68	37	79	80	82	84
8. ZIP/PHDf_Petcr_PHDf_HD_PRHP	67	80	54	66	83	65	78		53	47	71	73	45	39	48	74	42	62	63	62	62
9. ZIP/PHDf Pinra PHDf HD	82	74	72	83	73	82	70	69		66	53	52	67	52	73	55	54	53	52	53	52
10. ZIP/PHDf Poptr_PHDf HD_I	87	70	72	92	69	88	67	66	83		48	47	93	53	82	50	55	**49**	48	49	**49**
11. ZIP/PHDf_Poptr_PHDf HD_III	66	87	57	67	90	65	82	82	70	67		93	48	36	51	82	39	67	70	65	64
12. ZIP/PHDf_Poptr_PHDf_HD_II	65	88	56	66	90	65	81	**84**	69	66	96		46	35	**49**	83	38	67	69	64	64
13. ZIP/PHDf_Poptr_PHDf_HD_IV	86	70	72	91	69	88	67	66	82	97	67	66		52	82	48	54	47	47	48	47
14. ZIP/PHDf_Sacof PHDf_HD	71	60	82	73	57	69	55	57	70	72	56	55	72		57	39	92	35	35	33	33
15. ZIP/PHDf Vitvi PHDf HD I	89	71	73	90	70	88	67	**64**	84	91	67	66	89	73		51	58	48	48	**48**	**48**
16. ZIP/PHDf Vitvi PHDf HD II	67	88	58	69	91	68	84	84	73	69	90	93	67	57	69		42	65	67	65	64
17. ZIP/PHDf Zeama_PHDf HD	74	61	83	73	58	70	56	58	69	73	57	55	72	96	73	57		38	37	35	34
18. ZIP/PHDf Zeama_hoxla	71	84	59	71	85	70	88	79	74	72	83	82	71	61	70	85	62		94	74	76
19. ZIP/PH Df Zeama_hoxlb	70	84	58	69	85	68	89	79	74	70	84	83	70	60	70	85	60	98		75	78
20. ZIP/PHDf_Zeama_hox2a	67	80	54	67	83	67	89	76	70	67	80	79	66	54	66	82	56	87	87		92
21. ZIP/PHDf_Zeama_Hox2b	67	82	53	66	83	67	90	77	71	66	80	79	67	53	67	82	55	88	88	97	

A identidade da porcentagem também pode ser calculado entre o HD conservado da SEQ ID N°: 180 (como representado pela SEQ ID N°: 234) e o HD conservado dos polipeptídeos utéis na realização da invenção. A identidade da porcentagem no HD conservado entre a sequência de polipeptídeos úteis na realização de métodos da invenção que variam entre 25 % e 70 % da identidade de aminoácido, como mostrado na Tabela D4.

Tabela D4: Resultados MatGAT para a similaridade e identidade global no HD conservado entre uma sequência de polipeptídeos.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. HD_Arath_PHDf HD_PRHA		**34**	**54**	72	36	72	38	36	58	68	38	36	70	54	70	34	54	36	33	29	27
2. HD Arath PHDf-HD HAT3	56		**28**	32	54	34	40	58	30	32	54	58	32	28	32	58	28	42	39	40	39
3. HD_Orysa_PHDf HD	76	60		**68**	**42**	**58**	**42**	**36**	**46**	**64**	**38**	**40**	**66**	**94**	**62**	**42**	**92**	**38**	**35**	**35**	**33**
4. HD_Eucgr_PHDf HD	86	58	78		42	72	44	42	60	82	38	38	84	66	76	42	66	44	41	35	33
5. HD_Lotja_PHDf-HD	60	72	66	64		38	52	58	36	42	64	68	42	42	40	68	40	56	55	50	46
6. HD_Medtr_PHDf_HD	78	60	70	82	66		36	36	62	68	38	36	70	58	66	36	58	40	37	27	27
7. HD_Orysa_PHDf-HD_HAZ1	56	54	64	62	66	58		**44**	26	44	50	48	44	42	40	46	42	70	65	73	71
8. HD_Petcr_PHDf_HD_PRHP	58	74	58	56	66	58	58		36	42	54	58	42	36	38	62	36	50	45	40	39
9. HD Pinra PHDf HD	74	54	64	72	58	72	50	52		58	30	32	60	44	56	36	44	28	28	21	19
10. HD_Poptr PHDf HD_I	86	56	78	92	64	**84**	64	56	74		36	36	98	64	72	42	64	42	39	35	33
11. HD_Poptr_PHDf_HD_III	54	72	62	60	82	58	66	66	54	58		88	36	38	38	72	38	52	49	46	44
12. HD_Poptr_PHDf_HD_II	56	72	64	62	82	60	68	66	56	60	100		36	40	38	78	40	50	47	46	44
13. HD_Poptr_PHDf_HD_IV	88	56	80	94	64	86	**64**	56	76	98	58	60		66	72	42	66	42	39	35	33
14. HD_Sacof PHDf_HD	74	60	94	76	64	72	62	56	62	76	60	62	78		64	**42**	98	36	33	35	33
15. HD Vitvi PHDf HD I	82	60	80	86	70	80	60	58	74	88	62	64	88	82		38	64	40	37	29	27
16. HD Vitvi PHDf HD II	64	76	66	66	82	68	64	74	56	66	84	**84**	66	64	68		40	**48**	43	44	44

17. HD_Zeama_PHDf_HD	74	60	94	76	64	72	62	56	62	76	60	62	78	100	82	64		36	33	35	33
18. HD_Zeama_hoxla	56	60	62	60	66	60	78	60	52	62	62	64	62	58	60	68	56		86	60	56
19. HD_Zeama_hoxl b	57	63	61	61	69	61	80	63	51	63	67	69	63	57	61	71	55	92		59	53
20. HD_Zeama_hox2a	58	60	60	60	65	64	89	60	52	62	64	64	62	60	62	64	58	77	81		90
21. HD_Zeama_Hox2b	56	56	56	56	62	58	85	54	48	58	60	60	58	56	58	62	54	71	77	92	

Exemplo 32: A identificação dos domínios compreendidos nas sequências de polipeptídeos úteis na realização de métodos da invenção

O recurso integrado das famílias de proteínas, banco de dados de locais e domínios (InterPro) é uma interface integrada para os bancos de dados assinados comumente usados para as buscas com base no texto e sequência. O banco de dado InterPro combina este banco de dados, que usam as metodologias diferentes e variam os graus de informação biológica sobre as proteínas bem caracterizadas para derivar os sinais da proteína. A colaboração dos bancos de dados incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. Interpro é hospedado no European Bioinformatics Institute in the United Kingdom.

Os resultados da avaliação InterPro de uma sequência de polipeptídeo como representado pela SEQ ID N°: 180 são apresentados na Tabela D5.

Tabela D5: Resultados de avaliação InterPro de uma sequência de polipeptídeo como representado pela SEQ ID N°: 180

Número de acessão InterPro	nome do banco de dados integrado	número de acessão do banco de dados integrado	nome de acessão do banco de dados integrados	coordenadas de aminoácidos na SEQ IS N°:180
IPR0013556 Homeobox	ProDom	PD000010	PRH ARATH P48785	439-497
IPR0013556 Homeobox	PFAM	PF00046	Homeobox	439-495
IPR0013556 Homeobox	SMART	SM00389	HOX	438-500
IPR0013556 Homeobox	perfil	PS50071	Homeobox_2	436-496
IPR00195 dedo de Zn, tipo PHD	PFAM	PF00628	PHD	197-251
IPR00195 dedo de Zn, tipo PHD	SMART	SM00249	PHD	197-249

IPR00195 dedo de Zn, tipo PHD	perfil	PS50016	ZF PHD 2	195-251
IPRO12287 homeodomínio relacionado	Gene3D	G3DSA:1.10.10.60	nenhuma descrição	436-501
IPR013256 cromatina SPT2	SMART	SM00784	nenhuma descrição	78-154
IPR não integrado	PANTHER	PTHR19418	proteína Homeobox	397-414 431-590

Os resultados da avaliação InterPro claramente indentificam as características essenciais de um polipeptídeo PHDf-HD, isto é, um dedo de zinco PHDf e homeobox, como representado por exemplo respectivamente pelas entradas Pfam PF00628 e PF00046.

Os HD são conhecidos por ter resíduos canônicos. O HD dos polipeptídeos PHDf-HD é altamente divergente ainda em sequências nas posições que são quase invariáveis entre os homeodomínios. A sequência logo do HD dos polipeptídeos PHDf-HD da Tabela D1, é mostrado na Figura 15. os logos da sequência são uma representação gráfica por um alinhamento da sequência múltipla de ácido nucleico ou aminoácido. cada logo consiste de pilhas de símbolos, uma poilha para cada posição na sequência. A altura total da pilha indica uma conservação de sequência em que posição, enquanto a altura dos símbolos dentro da pilha indica a frequência relativa de cada amino ou ácido nucleico em que posição. No geral, um logo de sequência fornece uma descrição mais rica e precisa de; por exemplo, um local de ligação, do que seria uma sequência de consenso. O algorítmo (WebLogo) para produzir tal logo é disponível no servidor da University of California, Berkeley. O HD como representado pela SEQ ID N°: 234, e compreendido na SEQ ID N°: 180, é de acordo com um logotipo da sequência como representado na Figura 15. Os polipeptídeos úteis na realização de métodos de acordo como a invenção compreende um HD compreendido na logo de sequência como mostrado na Figura 15.

Exemplo 33: Predição das características da estrutura secundária de uma

sequência de polipeptídeos úteis na realização de métodos da invenção

Os espirais enrolados usualmente contém um padrão de resíduos de sete aminoácidos repetidos denominado repetições *heptad*. Os espirais enrolados são importantes para identificar para as interações proteína-proteína, tal como oligomerização de proteínas idênticas, of de proteínas da mesma família, ou de proteínas não relacionadas. Recentemente muitos progressos foram feitos na predição computacional do espiral enrolado dos dados de sequência. Muitos algorítmos bem conhecidos a uma pessoa habilitada na técnica estão disponíveis; as ferramentas ExPASy Proteomics. Um destes, COILS, é um programa que compara uma sequência 1 aos bancos de dados de espirais enrolados de dois paralelos filamentados de deriva-se de uma contagem de similaridade. Em comparação a esta contagem às distribuições de contagens nas proteínas de espiral enrolado e globular, o programa então calcula a probabilidade que uma sequência adotará de uma configuração do espiral enrolado

O polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180, tem dois domínios de espiral enrolado prodito no terminal N, com uma alta probabilidade, em todas as três janelas (14, 21 e 28) examinadas. Na Tabela D6, as coordenadas dos resíduos, resíduos, como as três janelas e valores de probabilidades correspondentes são mostrados. Na Figura 16, é a produção gráfica do algorítmo COILS no polipeptídeo como representado pela SEQ ID N°: 180, onde os dois espirais enrolados preditos em visível claridade na metade do terminal N do polipeptídeo, em todas as três janelas (como representado pelas tr~es linhas).

Tabela D6: Produção numérica do algorítmo COILS no polipeptídeo como representado pela SEQ ID N°: 180. As coordenadas dos resíduos (#), resíduos, as três janelas e valores de probabilidades correspondentes são mostrados. As probabilidades acima de 0,09 são mostrados em cinza.

#	*Resíduo*	Janela=l4	Prob	Janela=21	Prob	Janela=28	Probab.
87	P	d	0,001	*e*	0,001	E	0,001
88	T	e	0,063	f	0,274	F	0,099

89	R	f	0,063	g	0,567	G	0,099
90	R	g	0,063	a	0,567	A	0,099
91	K	b	0,066	b	0,642	B	0,188
92	H	c	0,066	c	0,642	C	0,188
93	K	d	0,069	d	0,642	D	0,591
94	Q	e	0,178	*e*	0,642	E	0,672
95	K	f	0,335	f	0,642	F	0,672
96	R	g	0,335	g	0,642	G	0,772
97	K	a	0,335	a	0,642	A	0,772
98	N	b	0,347	b	0,642	B	0,772
99	D	c	0,347	c	0,642	C	0,772
100	E	d	0,347	d	0,642	D	0,772
101	S	*e*	0,347	*e*	0,642	E	0,808
102	D	f	0,347	f	0,642	F	0,808
103	E	g	0,347	g	0,642	g	0,808
104	V	a	0,347	a	0,642	a	0,808
105	S	b	0,347	b	0,642	b	0,808
106	R	c	0,347	c	0,642	c	0,808
107	M	d	0,347	d	0,642	d	0,808
108	E	e	0,347	e	0,642	e	0,808
109	K	f	0,347	f	0,642	F	0,808
110	R	g	0,347	g	0,642	g	0,808
111	A	a	0,347	a	0,642	a	0,808
112	R	b	0,307	b	0,514	b	0,808
113	Y	c	0,066	c	0,503	c	0,808
114	L	d	0,262	d	0,503	d	0,808
115	L	e	0,262	e	0,503	e	0,808
116	I	f	0,262	f	0,503	F	0,808
117	K	g	0,262	g	0,503	g	0,808
118	I	a	0,262	a	0,503	a	0,808
119	K	b	0,262	b	0,503	b	0,808
120	Q	c	0,262	c	0,503	c	0,808
121	E	d	0,262	d	0,503	d	0,808
122	Q	e	0,262	e	0,503	e	0,808
123	N	f	0,262	f	0,503	F	0,808
124	L	g	0,262	g	0,434	g	0,808
125	L	a	0,262	a	0,434	a	0,808
126	D	b	0,262	b	0,434	b	0,808
127	A	c	0,262	c	0,304	c	0,808
128	Y	d	0,145	d	0,304	d	0,808
129	S	e	0,145	e	0,044	e	0,808
130	G	f	0,119	f	0,009	F	0,367
131	D	g	0,078	g	0,007	g	0,136
132	G	a	0,002	a	0,001	a	0,005
133	W	b	0	a	0,001	a	0,001
134	N	b	0,001	b	0,045	b	0,057
135	G	c	0,001	c	0,045	c	0,057
136	H	d	0,003	d	0,07	d	0,219
137	S	e	0,008	e	0,132	e	0,231
138	R	f	0,025	f	0,357	F	0,231
139	E	g	0,025	g	0,357	g	0,231
140	K	a	0,025	a	0,357	a	0,231
141	I	b	0,028	b	0,357	b	0,231
142	K	c	0,07	c	0,357	c	0,462
143	P	d	0,07	d	0,357	d	0,462
144	E	e	0,998	e	0,997	e	0,974
145	K	f	0,998	f	0,997	f	0,974
146	E	g	0,998	g	0,997	g	0,974

147	L	a	0,998	a	0,997	a	0,974
148	Q	b	0,998	b	0,997	b	0,974
149	R	c	0,998	c	0,997	c	0,974
150	A	d	0,998	d	0,997	d	0,974
151	K	e	0,998	e	0,997	e	0,974
152	K	f	0,998	f	0,997	f	0,974
153	Q	g	0,998	g	0,997	g	0,974
154	I	a	0,998	a	0,997	à	0,974
155	M	b	0,998	b	0,997	b	0,974
156	K	c	0,998	c	0,997	c	0,974
157	Y	d	0,998	d	0,997	d	0,974
158	K	e	0,989	e	0,997	e	0,974
159	I	f	0,896	f	0,997	f	0,974
160	A	g	0,486	g	0,997	g	0,974
161	I	a	0,409	a	0,997	a	0,974
162	R	b	0,274	b	0,997	b	0,974
163	D	c	0,255	c	0,997	c	0,974
164	V	d	0,095	d	0,997	d	0,974
165	I	e	0,025	e	0,871	e	0,974
166	H	f	0,025	f	0,622	f	0,974
167	Q	g	0,025	g	0,496	g	0,974
168	L	a	0,025	a	0,496	a	0,974
169	D	b	0,025	b	0,468	b	0,974
170	L	c	0,005	c	0,111	c	0,974
171	C	d	0,002	d	0,02	d	0,974
172	S	e	0,001	e	0,007	e	0,719
173	S	f	0,001	f	0,003	f	0,482
174	S	g	0,001	g	0,001	g	0,066
175	G	a	0	a	0	a	0,001

Um outro espiral enrolado é predito na metade do terminal C do polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180, com uma probabilidade inferior do que os dois domínios enrolados compreendidos na metade do terminal N do polipeptídeo.

Exemplo 34: Predição de localização subcelular de uma sequência de polipeptídeos úteis na realização de métodos da invenção

LOCtree é um algorítmo que pode predizer uma localização subcelular e propensão de ligação de DNA de proteínas de nenhuma membrana em plantas eucariotas e nenhuma planta tão bem quanto procariotas. LOCtree classifica as proteínas animais eucarióticas em uma das cinco classes subcelulares, enquanto as proteínas vegetais são classificadas em um a seis classes e proteínas procarióticas são classificadas em um a três classes.

Quando disponível, o LOCtree também diz respeito às predições com base nos seguintes: 1) Sinais de localização nuclear observados pelo algorítmo PredictNLS, 2) Localização inferida usando os motivos Prosite

e domínios Pfam observados na proteína, e 3) palavra chave SWISS-PROT associada com a proteína. A localização é inferida nos últimos dois casos usando o algorítmo LOCkey com base em entropia. O software é hospedado na University of Columbia, USA.

Motivo e palavra chave com base na predição da localização subcelular de um polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180, usando LOCkey:

Localização predita	Confiança	Predição alternativa	Palavra chave SWISS-PROT usada para determinar a localização
Nuclear	100	-	Homeobox, ligação de DNA, regulamento de transcrição, proteína nuclear, transcrição, dedo de zinco.

A predição de um sinal de localização nuclear (NLS) é feita usando o algorítmo PredictNLS, por exemplo. O algorítmo também é hospedado pelo servidor na University of Columbia, USA. Na Tabela abaixo, a predição de NLS no polipeptídeo da SEQ ID N°: 180 usando o algorítmo PredictNLS, é mostrado.

motivo generalizado ([KR]{3,5} entre 3 e 5 vezes Kou R)	acerto no polipeptídeo da SEQ ID N°:180	coordernadas no polipeptídeo da SEQ ID N°: 180
K[RK]{3,5}x{11,18}[RK]Kx{2,3}K	KRRRGSDAATGKSATGPTRRKHKQK	71-95
[KR]{4}x{20,24}K{1,4}xK	KRRRGSDAATGKSATGPTRRKHKQKRK	71-97

Na sequência de polipeptídeo abaixo (SEQ ID N°: 180), a posição do NLS predito é mostrado em negrito e sublinhado duas vezes:

MNTPEKKPLCYTSRRALQQRTESSSELISVSKRATRQNTPRKPDSPPKRTTRSSANLAKC
IENKHHSSPL**KRRRGSDAATGKSATGPTRRKHKQKRK**NDESDEVSRMEKRARYLLIKIKQ
EQNLLDAYSGDGWNGHSREKIKPEKELQRAKKQIMKYKIAIRDVIHQLDLCSSSGSKDDS
VIPPDGCHESVNPEHTICSRCKSHESFPDNNIIFCEGGCKLACHQKCLEPPFDKILPTTR
HGRLCKHCSSKMKILDAINAHLGTSFTVKCPSSDIFKEAAEHFNSDDGLGQDWLSEYSGD
EDYDPEENEASSSGEENKSADSNCSGSPLYSPNDDIPDFISADFNDAEGFCRESSNLGID
FGEDGLAEILTHQRPRRDVDYTQLNEQMFGEPIGNDEQSEDEDWGLNKRKKRRTGSTGVG
TNSVEGRSDVKSNKKAQPRRKLFRIPPAAVEVLRKAFAENELPARSVKENLSTELGISFE
KIDKWFKNTRCAALRDRKGESRYSGPSKRSRTSIEKAETSAKVDQMDNSCFLPLSEIINV
PTRLQKGLDKKPKSINSPPRPQDNETCLSPTDKTKEGTPPTIKPSITDSSQLMNNDIGTE
ETAVSWVDTWASDALHFLDVSDDEHFFDVIEKVCGLENRLQRLKENMLSSSSSTDNNVAA
ESGLQNEVVLVPAAELKDKAS

Muitos outros algorítmos podem ser usados para realizar tais análises, incluindo:

• ChloroP 1.1 hospedado no servidor da Technical University of Denmark;

• Protein Prowler Subcellular Localisation Predictor versão 1.2 hospedado no servidor do Institute for Molecular Bioscience, University of Queensland,. Brisbane, Australia;

• PENCE Proteome Analyst PA-GOSUB 2.5 hospedado no servidor da University of Alberta edmonton, Alberta, Canada;

• TMHMM, hospedado no servidor da Technical University of Denmark

Exemplo 35: Clonagem da sequência de ácido nucleico como representado pela SEQ ID N°: 179

A não ser outra maneira estado, as técnicas de DNA recombinante foram realizados de acordo com os protocolos padrões descritos em (Sambrook (2001) Molecular Cloning: a laboratory manual, 3° Edition Cold Spring Harbor Laboratory Press, CSH, New York) ou in Volumes 1 and 2 of Ausubel et at. (1994), Current Protocols in Molecular Biology, Current Protocols. Os materiais padrões e métodos para o trabalho molecular vegetal são descritos em Plant Molecular Biology Labfax (1993) by R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (UK) e Blackwell Scientific Publications (UK).

O gene de *Oryza sativa* PHDf-HD foi amplificado por PCR usando como modelo um banco de DNA de arroz sintetizado do mRNA extraído a partir de tecidos das plantas. O iniciador prm09687 (SEQ ID N°: 236; sentido,: 5'-ggggacaagtttgtacaaaaaagcaggcttaaacaatgaataccccagaaaa gaaa-3') e iniciador prm09688 SEQ ID N°: 237; reverso, complementar,: 5'-ggggaccactttgtacaagaaagctgggtgatgcaaggttaagatgcttt-3'), que incluem os locais AttB para a recombinação Gateway, foram usados pela amplificação

PCR. O PCR foi realizado usando a polimerase Hifi Taq de DNA em condições padrões. Um fragmento de PCR do comprimento esperado (incluindo locais attB) foi amplificado e purificado também usando os métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante o qual o fragmento PCR recombinado *in vivo* com o plasmídio pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada". O plasmídio pDONR201 foi adquirido de Invitrogen, como partes da tecnologia Gateway®.

Exemplo 36: Construção do vetor de expressão usando a sequência do ácido nucleico como representado pela SEQ ID N°: 179

O clone de entrada que compreende a SEQ ID N°: 179 foi subsequentemente usado em uma reação LR com um vetor de destino usado para a transformação da *Oryza sativa*. Este vetor contido como elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; um cassete de expressão marcador avaliável; e um cassete Gateway pretendido pela recombinação LR *in vivo* com uma sequência do ácido nucleico de interesse já clonado no clone de entrada. Um promotor GOS2 de arroz (SEQ ID N°: 235) para a expressão constitutiva foi localizada a montante deste cassete Gateway.

Após a etapa de recombinação LR, o vetor de expressão resultante pGOS2::PHDf-HD e (Figura 18) foi transformado na cepa da *Agrobacterium* LBA4044 de acordo com os métodos bem conhecidos na técnica.

Exemplo 37: Transformação da planta

Transformação de arroz

A *Agrobacterium* contendo os vetores de expressão foram usados independentemente para transformar as plantas *Oryza sativa*. As sementes secas maduras de japonica cultiva de arroz Nipponbare foram descascadas. A esterilização foi realizada pela incubação em um minuto em

70 % de etanol, seguido por 30 minutos em 0,2 % de $HgCl_2$, seguido por 6 vezes de 15 minutos lavada com água destilada estéril. As sementes estéreis foram então germinadas em um meio contendo 2,4-D (meio de indução de calos). Após a incubação no escuro por quatro semanas calos derivados de escutelo embriogênico foram taxados e propagados no mesmo meio. Após duas semanas, os calos foram multiplicados ou propagados pela subcultura no mesmo meio por outras 2 semanas. Os pedaços de calos embriogênicos foram sub-culturados em meio fresco 3 dias antes da co-cultivação (para intensificar a atividade de divisão celular).

A cepa de *Agrobacterium* LBA4404 contendo cada vetor de expressão individual foi usado independentemente para a co-cultivação. A *Agrobacterium* foi inoculada no meio AB com os antibióticos apropriados e cultivados por 3 dias a 28° C. As bactérias foram então coletadas e submetidas à suspensão no meio de co-cultivação líquido a uma densidade (OD_{600}) de cerca de 1. A suspensão foi então transferida a um disco Petri e os calos sumetidos à imersão na suspensão por 15 minutos. Os tecidos dos calos foram então manchados secos em um papel de filtro e transferidos para solidificar-se, meio de co-cultivação e incubados por 3 dias no escuro a 25° C. Os calos co-cultivados foram desenvolvidos em meio contendo 2,4-D por 4 semanas no escuro a 28° C na presença de um agente de seleção. Durante este período, os calos resistentes rapidamente ao desenvolvimento foram ilhas desenvolvidas. Após a transferência deste material material a um meio de regeneração e incubação na luz, o potencial embriogênico foi liberado e os brotos desenvolvidos nas próximas quatro ou cinco semanas. Os brotos foram taxados dos calos e incubados por 2 a 3 semanas em um meio contendo auxina de que estes foram transferidos ao solo. Os brotos endurecidos foram desenvolvidos sob alta umidade e dias curtos em uma estufa.

Aproximadamente 35 transformantes de arroz TO independentes foram gerados para cada construção. Os primeiros

transformantes foram transferidos a partir de uma câmera de cultura de tecido de uma estufa. Após uma análise PCR quantitativa para verificar o número de cópias da inserção de T-DNA, apenas uma simples cópia da planta transgênica que exibe a tolerância ao agente de seleção foi mantida para a coleta da semente T1. As sementes foram então coletadas três a cinco meses após a transplantação. O método produzido de transformantes de local simples em uma taxa acima de 50 % (Aldemita and Hodges1996, Chan *et al.* 1993, Hiei *et al.* 1994).

Exemplo 38: Procedimento de avaliação fenotípica

38.1 Apresentação da avaliação

Aproximadamente 35 transformantes de arroz TO independentes foram gerados. Os primeiros transformantes foram transferidos a partir de uma câmera de cultura de tecido de uma estufa para o desenvolvimento e coletado da semente T1. Seis eventos, de que uma progênie T1 segregada 3:1 para a presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 mudas T1 contendo o transgene (hetero e homozigotos) e aproximadamente 10 mudas T1 perderam o transgene (nulizigotos) foram selecionados pela expressão marcadora visual do monitoramento. As plantas transgênicas e os nulizigotos correspondentes foram desenvolvidos lado a lado em posições aleatórias. As condições da estufa foram de dias curtos (12 horas de luz), 28° C na luz e 22° C no escuro, e uma umidade relativa de 70 %. O desenvolvimento das plantas sob as condições de não tensão são

fornecidos com água em intervalos regulares para garantir que a água e nutrientes não são limitantes para satisfazer a necessidade das plantas para completar o crescimento e desenvolvimento.

A partir do estágio de semeadura até o estágio de maturidade das plantas foram passados diversas vezes através de um gabinete de imagem digital. Em cada período de tempo imagens digitais (2048x1536 píxeis, 16

milhões de cores) foram tiradas de cada planta em pelo menos 6 ângulos diferentes.

38.2 Análise estatística: teste F

Dois fatores ANOVA (análise de variantes) foi usado como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformados com o gene da presente invenção. O teste F foi realizado para checar por um efeito do gene em todos os eventos de transformação e para verificar por um total do efeito do gene, também conhecido como um efeito do gene global. O limiar para a significância por um efeito verdadeiro do gene global foi apresentado em uma 5 % do nível de probabilidade para o teste F. Um ponto do valor do teste F significante por um efeito do gene, significando que este não é apenas uma mera presença ou posição do gene que está causando as diferenças no fenotipo.

38.3 Parâmetros medidos

Medição do parâmetro relacionado à biomassa

A partir do estágio de semeadura até o estágio de maturidade das plantas foram passados diversas vezes através de um gabinete de imagem digital. Em cada período de tempo imagens digitais (2048x1536 píxeis, 16 milhões de cores) foram tiradas de cada planta em pelo menos 6 ângulos diferentes.

Uma área da planta acima do solo (ou biomassa da folha) foi determinado pela contagem do número total de píxeis nas imagens digitais a partir das partes de plantas acima do solo discriminadas a partir do fundamento. Este valor foi medido pelas figuras retiradas do mesmo ponto de tempo a partir de ângulos diferentes e foi convertido a um valor de superfície física expressada em quadrado mm pela calibração. Os experimentos mostram que à área da planta acima do solo medida desta maneira correlaciona com a biomassa das partes das plantas trituradas acima. À área triturada acima é a

área medida no ponto de tempo em que a planta tem atingido esta biomassa máxima da folha. O vigor precoce é à área das planta (mudas) acima do solo três semanas pós-germinação. O aumento na biomassa da raiz é expressado como um aumento no total da biomassa da raiz (medido como biomassa máxima de raízes observadas durante a vida toda de uma planta); ou como um aumento na raiz/índice de broto (medido como a razão entre massa de raiz e massa de broto no período de desenvolvimento ativo da raiz e broto).

Medições dos parâmetros relacionados à semente

As panículas primárias maduras foram coletadas, contadas, ensacadas, rotuladas com código de barra e então secadas por três dias em um forno a 37° C. As panículas foram então limiadas e todas as sementes foram coletadas e contadas. As cascas enchidas foram separadas a partir de um vazio usando um dispositivo de sopro de ar. As cascas vazias foram descartadas e a fração remanescente foi contada novamente. As cascas enchidas foram pesadas em uma balança analítica. O número de sementes enchidas foi determinado pela contagem do número de cascas enchidas que permanece depois da etapa de separação. O peso total da semente por planta foi medida pela pesagem total das cascas enchidas coletadas a partir de uma planta. O número total de semente por planta foi medido pela contagem do número de cascas coletadas a partir da planta. Peso de núcleo milhar (TKW) é extrapolado do número de sementes enchidas contadas e seu peso total. O índice de colheita (HI) na presente invenção é definida como a razão entre o peso total da semente por planta e à área triturada acima (mm^2), multiplicada por um fator 10^6. O número total de flores por panícula como definido na presente invenção é a razão entre o número total de sementes e o número de panículas primárias maduras. A taxa de abastecimento da semente como definido na presente invenção é a proporção (expressado como uma %) do número de sementes enchidas no número total de sementes (ou florzinhas).

Avaliação da disponibilidade do nutriente reduzido (nitrogênio)

As plantas de arroz das sementes T2 foram desenvolvidas no solo em potes sob condições normais exceto para a solução do nutriente. Os potes foram submetidos à água a partir da transplantação do amadurecimento com uma solução de nutriente específico contendo teor de nitrogênio N reduzido (N), usualmente entre 7 a 8 vezes menor. O resto do cultivo (amadurecimento da planta, coleta da semente) foi a mesma como para as plantas não desenvolvidas sob a tensão abiótica. Os parâmetros de produção e desenvolvimento foram registrados como detalhado para o desenvolvimento sob as condições normais.

Exemplo 39: Resultados da avaliação fenotípica das plantas transgênicas de arroz

Os resultados da avaliação das plantas transgênicas de arroz que expressam uma sequência do ácido nucleico que codifica um polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180, sob o controle do promotor GOS2 para a expressão constitutiva, são apresentadas abaixo.

Há um aumento significante no número de panículas, na produção total de semente por planta, no número total de sementes enchidas, no número total de sementes, nos peso de núcleo milhar (TKW), e no índice de colheita das plantas transgênicas comparados aos nulizigotos correspondentes (controles), como mostrado na Tabela D7.

Tabela D7: Resultados da avaliação das plantas transgênicas de arroz que expressam uma sequência do ácido nucleico que codifica um polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180, sob o controle do promotor GOS2 para a expressão constitutiva.

	Média da % do aumento em 3 eventos na geração T1
Número das primeiras panículas	22 %
Produção da semente total por planta	30 %
Número total de sementes enchidas	24 %
Número total de sementes	24 %
TKW	4 %
Índice de colheita	17 %

Exemplo 40: Resultados da avaliação fenotípica das plantas transgênicas do

desenvolvimento de arroz sob a disponibilidade do nutriente reduzido

Os resultados da avaliação das plantas transgênicas de arroz que expressam uma sequência do ácido nucleico que codifica um polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180, sob o controle do promotor GOS2 para a expressão constitutiva, e desenvolvimento sob a disponibilidade do nutriente reduzido, são apresentadas abaixo.

Há um aumento significante na biomassa, no vigor emergido, na produção total de semente por planta, no número total de sementes enchidas, na taxa de enchimento da semente, no número total de sementes, nos peso de núcleo milhar (TKW), e no índice de colheita das plantas transgênicas comparados aos nulizigotos correspondentes (controls), como mostrado na Tabela D8.

Tabela D8: Resultados da avaliação das plantas transgênicas de arroz que expressam uma sequência do ácido nucleico que codifica um polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180, sob o controle do promotor GOS2 para a expressão constitutiva, e desenvolvimento sob a disponibilidade do nutriente reduzido.

	% de aumento médio total (6 eventos na geração T2)
Biomassa	8 %
Vigor emergente	20 %
Produção de semente total por planta	17 %
Número total de sementes enchidas	12 %
Taxa de enchimentos da sementes	3 %
Número total de sementes	9 %

Exemplo 41: Exemplos de transformação de outras lavouras

A transformação de milho, trigo, soja, colza/canola, alfafa e algodão foi realizado como resumido no Exemplo 7

Exemplo 42: Exemplos de avaliações de tensão abiótica

Avaliação da aridez

As plantas foram selecionadas de número de eventos e são desenvolvidas no solo em potes sob condições normais até estes aproximarem do estágio superior. Estas são então transferidos à seção "seca" onde a

irrigação é retida. As sondas de umidades são inseridas em potes escolhidos aleatoriamente para monitorar o teor de água no solo (SWC). Quando o SWC está abaixo de certos limiares, as plantas são automaticamente novamente submetidas à água continuamente até um nível normal ser atingido novamente. As plantas são então re-transferidas às condições normais. O resto do cultivo (amadurecimento da planta, coleta da semente) é o mesmo como para as plantas não desenvolvidas sob as condições de tensão abiótica. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais.

Avaliação da tensão do sal

As plantas são desenvolvidas em um substrato feito de fibras de coco e argex (razão de 3 para 1). Uma solução de nutriente normal é usado durante as primeiras duas semanas após a transplantação das plantinhas nas estufas. Após as primeiras duas semanas, 25 mM de sal (NaCl) é adicionado à solução de nutriente, até as plantas serem coletadas. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais.

VIII. Proteína semelhante à bHLH11 (hélice 11-arco-hélice básico)

Exemplo 43: A identificação das sequências relacionadas à sequência de ácido nucleico semelhante à bHLH11 usado nos métodos da invenção

As sequências (cDNA eSTs de comprimento total ou genômico) relacionadas a uma sequência do ácido nucleico usado nos métodos da presente invenção onde identificados entre aqueles mantidos no banco de dados Entrez Nucleotides at the National Center for Biotechnology Information (NCBI) usando as ferramentas de sequência de banco de dados, tal como a ferramenta de alinhamento local básica (BLAST) (Altschul et al. (1990) J. Mot. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa é usado para observar as regiões da similaridade local entre as sequências em comparação ácido nucleico ou sequências de

polipeptídeos de um banco de dados da sequência e pelo cálculo da significância estatística de pontos. Por exemplo, o polipeptídeo codificado pelo ácido nucleico usado na presente invenção foi usado pelo algorítmo TBLASTN, com ajustes default e o filtro para ignorar a baixa complexidade da sequência apresentada. A produção da análise foi examinada pela comparação em pares, e classificada de acordo com as contagens de probabilidade (valor E), onde a contagem reflete a probabilidade que um alinhamento particular ocorra pela possibilidade (o valor E inferior, quanto mais significante o acerto). Além disso aos valores E, as comparações também foram contadas pela identidade de porcentagem. A identidade da porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácido nucleico comparado (ou polipeptídeo) em um comprimento particular. Em alguns exemplos, os parâmetros default podem ser ajustados para modificar a estringência da busca. Por exemplo o valor e pode ser aumentado para mostrar os menores pontos estringentes. Desta maneira, os pontos curtos próximos exatos podem ser identificados. Tabela E1 fornece uma lista das sequências de ácidos nucleicos relacionadas a uma sequência do ácido nucleico usado nos métodos da presente invenção.

Tabela E1: Exemplos de polipeptídeo semelhante aos bHLH11:

fonte da planta	**ácido nucleico da SEQ ID N°:**	**proteína da SEQ ID N°:.**
Triticum aestivum	244	245
Allium cepa	258	327
Arabidopsis thaliana	259	328
Arabidopsis thaliana	260	329
Arabidopsis thaliana	261	330
Arabidopsis thaliana	262	331
Arabidopsis thaliana	263	332
Aquilegia vulgaris	264	333
Aquilegia vulgaris	265	334
Brassica napus	266	335
Citrus clementina	267	336

Curcuma longa	268	337
Citrus paridisi hybrid	269	338
Citrus sinensis	270	339
Eucalyptus grandis	271	340
Eucalyptus grandis	272	341
Gossypium hirsutum	273	342
Gossypium hirsutum	274	343
Gossypium hirsutum	275	344
Gossypium hirsutum	276	345
Gossypium hirsutum	277	346
Gossypium hirsutum	278	347
Glycine max	279	348
Glycine max	280	349
Glycine max	281	350
Glycine max	282	351
Gossypium raimondii	283	352
Helianthus petiolaris	284	353
Hordeum vulgare	285	354
Lactuca perennis	286	355
Nicotiana benthamiana	287	356
Nicotiana benthamiana	288	357
Nicotiana benthamiana	289	358
Nicotiana benthamiana	290	359
Nicotiana tabacum	291	360
Oryza sativa	292	361
Oryza sativa	293	362
Oryza sativa	294	363
Oryza sativa	295	364
Oryza sativa	296	365
Oryza sativa	297	366
Oryza sativa	298	367
Picea abies	299	368
Populus deltoides	300	369
Pinus radiata	301	370
Picea sitchensis	302	371

Pinus taeda	303	372
Populus trichocarpa	304	373
Populus trichocarpa	305	374
Populus trichocarpa	306	375
Populus trichocarpa	307	376
Populus trichocarpa	308	377
Poncirus trifoliata	309	378
Ricinus communis	310	379
Ricinus communis	311	380
Sorghum bicolor	312	381
Solanum lycopersicum	313	382
Solanum lycopersicum	314	383
Solanum lycopersicum	315	384
Solanum tuberosum	316	385
Solanum tuberosum	317	386
Solanum tuberosum	318	387
Solanum tuberosum	319	388
Triticum aestivum	320	389
Vitis vinifera	321	390
Vitis vinifera	322	391
Vitis vinifera	323	392
Vitis vinifera	324	393
Zea mays	325	394
Zea mays	326	395

Em alguns exemplos, as sequências relacionadas tem sido como tentativa reunida e publicada descobertas pelas instituições de busca, tal como The Institute for Genomic Research (TIGR). Os bancos de dados Eukaryotic Gene Orthologs (EGO) podem ser usados para identificar tais sequências relacionadas pela busca de palavra chave ou pelo uso de algorítmo BLAST com o ácido nucleico ou sequência de polipeptídeo de interesse.

Exemplo 44: Alinhamento das polipeptídeo semelhante às sequências bHLH11

O alinhamento das sequências de polipeptídeos foi realizado

usando o programa AlignX a partir do vetor NTI (Invitrogen) que é baseado no algorítmo Clustal W popular do alinhamento progressivo (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31:3497-3500). Os valores default são pela penalidade aberta de fenda de 10, pela penalidade de extensão de fenda de 0,1 e a matriz do peso selecionada é Blosum 62 (se os polipeptídeos são alinhados). Uma edição menor do manual foi feita ainda para otimizar o alinhamento. A conservação da sequência entre os polipeptídeos semelhante aos bHLH11 é essencialmente no domínio bHLH do domínio do terminal C dos polipeptídeos, o domínio do terminal N usualmente sendo mais variável no comprimento da sequência e composição. O polipeptídeo semelhante aos bHLH11 são alinhados na Figura 21.

Uma árvore filogenética de polipeptídeo semelhante aos bHLH11 (Figura 22) foi construído usando "CLUSTALX", é uma árvore próxima foi calculada. O cladograma circular foi tirado usando Dendroscope (Huson et al., 2007).

Exemplo 45: Cálculo da identidade global da porcentagem entre as sequências de polipeptídeos úteis na realização de métodos da invenção

As porcentagens globais de similaridade e identidade entre as sequências de polipeptídeos de comprimento total úteis na realização de métodos da invenção foram determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Ferramenta de alinhamento global de matriz) (BMC Bioinformatics. 2003 4:29. MatGAT: uma aplicação que gera as matrizes de similaridade/identidade usando as sequências de DNA ou proteínas. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). O software MatGAT gera as matrizes de similaridade/identidade para sequência de proteínas ou DNA sem necessariamento o pré-alinhamento dos dados. O programa realiza uma série de alinhamento em pares usando o algorítmo de alinhamento global Myers

and Miller (com uma penalidade de abertura de fenda de 12, e a penalidade de extensão de fenda de 2), calcula a similaridade e identidade usando por exemplo Blosum 62 (para os polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade da sequência é mostrada na metada da parte inferior da linha de divisão e a identidade de sequência é mostrada na metada da parte superior da linha de divisão diagonal.

Parâmetros usados na comparação foram:

Matriz de contagem:	Blosum62
Primeira fenda:	12
Extensão da fenda:	2

Os resultados da análise do software são mostrados na Tabela E2 pela similiaridade e identidade global no comprimento total de uma sequência de polipeptídeos. A identidade da porcentagem é dada acima em diagonal e a porcetagem da similaridade é dada abaixo em diagonal. A SEQ ID N°: 245 é representada como TabHLH11.

A identidade da porcentagem entre o polipeptídeo semelhante às sequências bHLH11 úteis na realização de métodos da invenção podem ser tão baixo quanto 20 % da identidade de aminoácido comparado a SEQ ID N°: 245. A identidade é entretanto muito mais alta quando os domínios HLH são comparados (Tabela E3).

Tabela E2: Resultados MatGAT para a similaridade e identidade global no comprimento total da sequência de polipeptídeo semelhante à bHLH11.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. TC3698		51,1	53,5	15,9	39,5	37,1	49,4	19,5	53,8	45,8	36,3	53,7	30,6	52,0	46,5	51,0	52,2	51,7	51,2	34,6
2. AT1 G03040	63,2		81,0	16,3	37,8	36,3	52,4	19,9	72,3	40,5	35,9	64,6	31,1	50,7	59,5	45,9	58,4	68,1	45,3	37,0
3. AT4G02590	64,2	89,0		17,0	37,0	35,1	55,2	18,8	82,4	43,0	35,4	69,5	33,8	53,5	64,6	47,6	62,6	72,6	47,0	36,4
4. AT2G24260	21,7	22,0	22,3		27,7	19,7	16,4	12,7	15,6	17,0	19,0	15,8	23,0	15,8	15,8	16,8	15,9	16,8	15,4	17,2
5. AT4G30980	53,2	53,9	52,6	30,6		46,4	36,0	22,0	35,5	38,7	40,6	35,1	41,3	36,3	32,6	36,2	37,7	36,1	37,4	37,2
6. AT5G58010	53,9	55,0	52,6	23,8	58,1		32,6	24,2	34,0	35,4	37,1	36,0	35,7	31,6	32,3	35,5	38,8	36,0	38,1	39,2
7. TC15501	58,4	64,1	65,5	22,1	49,9	48,1		17,5	52,8	47,0	37,0	56,0	35,1	57,8	59,2	52,8	52,4	60,9	51,0	31,4
8. TC19278	31,9	31,1	30,0	15,6	33,2	35,0	26,8		18,8	21,1	16,4	19,5	18,9	17,8	17,6	17,5	23,0	20,2	19,4	27,6
9. TC10015_part	64,9	80,5	85,8	19,9	50,6	52,2	60,4	31,7		41,9	35,6	72,5	30,8	52,5	59,7	47,1	61,4	68,0	51,4	35,2
10. DY268946	56,8	55,9	59,1	23,8	52,5	46,4	61,0	31,6	53,0		34,5	42,7	32,9	52,7	43,2	50,8	46,9	45,0	57,6	35,6
11. TA2544	47,3	46,8	47,8	26,1	49,3	46,3	49,0	23,9	45,6	47,6		35,0	44,9	38,4	35,9	36,4	33,4	36,1	34,6	31,1
12. TA3392	68,1	76,8	79,4	21,7	51,9	52,2	63,0	32,7	84,0	56,2	43,7		30,9	53,0	68,7	49,4	63,1	77,0	52,2	36,8
13. TA12416	38,0	38,2	40,7	30,4	48,4	43,3	46,3	25,8	37,3	43,7	54,6	40,3		32,2	32,4	33,1	29,7	33,3	32,1	31,3
14. WO051050seqID77	61,3	63,8	66,3	20,8	51,3	50,9	66,4	28,1	61,9	65,8	45,4	61,3	39,4		53,2	58,7	50,0	54,6	58,1	31,5
15. WO051050seqID1671	56,9	70,5	74,9	22,3	46,5	47,1	70,9	27,5	68,5	59,0	47,3	75,4	42,0	64,5		46,8	58,7	71,0	44,8	33,5
16. TA55042	62,5	59,5	62,8	22,4	51,7	52,0	65,8	27,8	58,0	67,0	48,8	58,9	42,4	69,2	62,7		48,1	48,6	58,1	33,8
17. TC207545	60,7	66,6	68,1	20,1	51,0	51,5	57,0	38,5	71,6	53,9	42,9	74,7	36,7	56,9	64,5	56,8		64,7	56,0	37,4
18. TA13791	63,7	80,5	82,6	22,1	52,6	54,8	68,1	31,7	78,2	59,4	48,3	83,5	43,1	65,0	78,0	63,1	71,6		50,6	35,4
19. CO123623	61,4	57,3	58,7	19,3	49,7	53,9	58,1	32,8	62,6	64,3	42,9	65,1	38,2	65,6	55,2	65,6	64,9	61,7		36,8
20. DT543504	52,0	55,2	54,8	22,3	51,3	52,6	47,3	37,9	49,7	50,4	42,7	55,2	39,7	51,9	51,4	50,5	52,6	54,9	50,0	
21. TC60118	63,4	80,2	82,3	22,0	52,3	54,5	67,8	31,7	77,9	59,1	48,0	83,2	42,9	64,7	77,7	62,8	71,3	99,7	60,7	55,2
22. TC60119	65,1	78,9	82,9	21,1	51,9	53,3	67,0	30,6	77,3	58,6	47,6	82,6	40,5	65,6	76,6	60,7	71,1	91,1	58,6	53,6
23. TC61833	63,6	78,4	81,9	21,9	54,2	55,4	66,1	30,2	77,0	56,8	47,6	82,3	40,7	65,0	76,9	62,5	71,1	90,8	56,4	54,9
24. TC67603	40,3	43,3	42,4	28,8	48,7	45,9	47,5	26,9	40,7	45,7	56,7	42,9	67,0	43,1	42,9	45,7	38,6	42,9	40,0	41,2
25. TC229602	56,8	60,3	61,0	18,9	45,8	49,2	50,7	45,2	63,3	49,9	38,0	65,1	33,5	52,5	57,5	52,3	81,1	65,3	61,9	46,1
26. TC205173	50,4	51,0	48,7	25,3	58,5	54,6	53,3	34,3	47,2	52,5	52,2	47,5	54,6	51,3	48,0	54,6	49,0	51,0	49,6	49,9
27. TA1140	63,4	64,1	64,5	21,5	55,2	53,9	64,7	31,0	62,7	55,1	45,1	63,1	42,2	65,3	60,1	64,0	61,1	66,0	62,4	53,9
28. TC140470	37,6	38,5	38,0	26,6	35,8	36,4	41,2	22,2	34,9	37,0	43,7	37,4	48,0	38,0	40,3	40,3	33,5	37,8	32,8	43,9
29. TA3490_part	67,6	64,6	64,2	22,0	53,2	54,5	62,1	32,4	65,2	58,0	45,6	67,9	41,4	63,4	59,5	63,4	63,1	68,6	63,1	52,6
30. CK293938	67,6	64,2	65,2	20,6	53,9	54,2	61,5	32,1	65,2	60,0	45,6	66,2	41,4	68,4	60,1	64,0	59,8	66,0	64,9	53,6
31. TC8633	64,6	78,8	79,0	22,5	53,9	56,3	68,1	31,8	74,5	56,5	45,1	75,5	41,8	65,3	71,1	62,5	69,9	80,9	59,3	54,9
32. TC7102	39,7	42,4	42,9	30,0	49,2	44,2	46,3	25,1	39,7	44,7	56,9	40,6	69,7	43,6	44,2	43,6	38,1	44,2	39,3	41,3

33. TC7103	45,0	46,0	46,3	26,7	53,2	51,2	48,8	27,9	45,0	46,5	57,1	45,2	65,0	46,3	50,4	49,4	42,9	49,1	41,6	45,5
34. TC12771	66,1	64,9	65,5	21,2	54,8	54,4	61,5	30,9	64,4	59,7	45,6	65,1	40,9	67,2	59,8	64,0	59,7	66,0	65,8	54,9
35. 0s03g58330	67,7	61,9	61,3	21,3	53,9	54,2	58,7	32,7	62,6	55,9	45,4	61,6	39,2	62,2	53,8	61,9	59,2	63,7	62,9	51,3
36. 0s06g08500	43,6	43,6	45,0	23,6	45,0	48,2	47,7	28,6	42,0	43,3	51,5	44,1	46,5	44,1	45,2	46,6	40,3	43,3	39,0	43,3
37. 0s06g09370	37,7	37,4	38,1	26,8	35,1	34,1	40,0	22,2	34,1	40,6	44,1	38,5	44,4	38,5	40,6	42,7	33,1	38,9	36,2	40,8
38. 0s02g55250	40,6	42,0	41,5	28,5	43,5	44,4	45,1	25,2	39,2	46,0	55,8	40,8	56,3	44,0	41,7	46,3	38,1	41,5	39,0	43,5
39. 0s02g35660	35,7	35,1	36,6	29,2	36,5	34,4	39,3	22,9	33,6	38,5	43,1	34,9	45,4	36,6	39,5	37,0	32,3	37,4	34,0	39,7
40. 0s07g08440	70,3	60,3	65,5	20,8	52,3	56,9	58,4	33,8	63,8	57,1	44,6	63,8	38,8	62,5	57,8	61,9	58,4	63,7	65,9	51,0
41. 0s09g25040	37,3	35,7	38,5	30,0	40,1	38,5	41,5	24,4	36,1	38,7	54,3	37,7	60,7	39,3	41,5	39,3	33,9	38,3	35,7	37,7
42. Pt_scaff_II.416	62,7	79,5	81,9	22,4	53,5	53,7	66,7	31,3	78,7	58,3	47,3	84,3	42,0	65,6	78,3	62,2	73,3	89,4	61,0	55,2
43. Pt_scaff_70.65	60,7	62,2	65,9	21,9	52,3	49,2	65,8	27,5	60,7	69,9	44,9	60,7	40,3	71,9	64,5	74,6	56,5	63,4	66,2	50,8
44. Pt_scaff_XII1.403	59,9	62,0	62,3	21,2	50,6	49,7	66,1	28,1	58,7	67,0	45,4	59,0	42,2	69,2	65,9	73,1	55,4	64,4	64,7	49,7
45. Pt_scaff28.86	39,5	39,3	40,8	27,2	39,3	38,8	42,5	29,4	36,2	43,4	45,6	39,9	48,0	40,1	42,1	42,3	36,0	39,5	36,8	50,2
46. Pt_TC63334	65,0	79,8	81,6	22,1	54,8	52,3	68,7	31,3	78,7	58,0	47,6	84,3	42,6	65,0	77,2	62,5	73,7	89,1	62,0	53,6
47. TA14134	46,2	49,8	50,8	24,7	50,5	50,2	49,3	37,5	43,7	52,5	44,6	45,2	41,2	50,2	47,1	51,4	45,5	50,2	46,8	50,8
48. TA5414	38,0	38,5	41,2	30,1	48,5	42,7	45,9	25,9	37,8	42,9	55,8	41,2	98,9	41,2	42,1	42,7	37,0	42,9	39,1	39,7
49. TA3263	65,3	79,5	81,6	22,3	54,8	54,0	68,4	31,7	79,0	58,3	48,3	84,7	42,6	65,0	77,2	62,8	73,7	89,4	62,3	54,2
50. TA2825	65,2	80,1	83,9	22,0	53,2	54,2	66,7	32,1	80,3	55,9	47,3	86,0	41,6	66,6	79,5	63,7	74,6	90,4	60,9	55,9
51. TA1616	61,0	65,3	66,3	22,3	52,1	50,3	70,4	28,8	60,4	70,4	47,8	62,9	41,6	74,2	63,9	73,4	57,7	66,6	68,4	54,0
52. TA21665	66,2	65,2	63,9	21,3	52,3	53,2	60,4	31,1	63,5	58,3	44,4	67,6	42,9	64,4	59,5	61,0	60,1	66,0	62,5	52,3
53. TC172581	64,5	63,8	67,1	21,1	52,9	53,9	61,0	30,6	62,2	60,0	46,6	66,1	41,8	66,3	61,3	63,4	59,5	67,1	62,2	54,2
54. AK247217	64,6	80,5	80,3	22,9	52,3	52,0	70,1	31,5	75,2	56,2	45,1	78,1	42,0	67,8	72,5	64,4	70,5	82,8	60,6	52,6
55. TC104646	69,5	61,6	61,9	20,7	51,0	53,5	58,7	35,0	67,3	55,7	42,0	63,3	39,2	60,3	56,1	59,8	63,5	63,0	65,3	49,7
56. DV982110	59,3	55,6	53,5	22,6	55,8	57,9	51,0	38,0	55,8	51,3	44,9	61,6	41,6	53,4	49,1	51,7	56,9	58,1	62,0	56,2
57. WO051050segID516	40,2	42,5	42,3	28,0	44,1	42,3	46,4	27,0	39,3	46,9	53,3	42,5	52,5	44,1	43,6	43,8	37,8	43,2	39,5	41,9
58. TC68930	48,3	49,2	49,5	25,9	49,5	53,5	47,0	38,4	45,6	51,3	47,3	47,4	42,0	50,5	46,2	52,0	48,3	50,2	45,0	50,5
59. TA30646	64,6	79,8	80,0	22,1	52,6	52,3	68,7	31,5	75,8	55,7	44,6	78,1	41,2	67,2	72,8	64,4	70,2	83,5	60,6	54,2
60. TA28621	65,8	64,1	67,1	21,3	53,2	52,3	61,5	30,3	62,8	60,9	46,8	67,1	41,6	65,6	61,8	64,4	59,5	68,1	63,5	54,2
61. TA34455	40,3	41,6	41,2	29,5	49,7	43,4	45,9	23,0	38,0	45,6	59,3	41,4	67,0	43,4	43,8	44,7	37,1	42,5	36,9	41,2
62. TA37666	45,5	46,6	48,7	27,4	53,9	48,2	51,8	28,3	43,7	50,3	60,7	47,1	63,1	48,4	48,4	49,0	42,1	49,0	45,3	44,5
63. TC253044	56,6	54,0	51,6	19,7	50,3	53,9	49,3	36,9	53,4	45,8	42,9	55,5	39,0	50,9	48,8	51,1	53,4	53,1	52,8	59,8
64. TabHLH11	71,2	60,9	61,0	19,9	50,6	56,6	57,0	34,2	66,2	55,9	42,7	66,2	38,8	59,7	55,8	60,1	61,9	62,7	66,9	50,3
65. TA46156	67,5	76,8	78,4	22,3	54,8	55,2	66,4	33,6	78,9	58,6	45,9	84,1	40,5	64,7	74,6	64,0	74,4	85,5	64,4	56,2
66. GSVIVT00017237001	43,5	42,1	43,1	29,4	48,1	45,2	47,4	27,5	40,4	47,8	59,1	40,4	69,5	44,0	45,5	45,7	40,9	45,9	39,7	43,8
67. TA46194	66,1	69,0	71,9	22,3	52,1	53,4	70,9	29,7	67,4	63,5	49,0	69,0	41,8	72,5	68,8	68,3	64,5	73,5	66,8	53,4
68. GSVIVT00016367001	35,7	37,7	36,5	24,5	38,1	36,9	41,6	24,1	34,4	41,9	44,5	35,5	46,2	37,7	40,8	38,6	32,8	37,1	34,6	45,4
69. TA139285	70,2	59,9	61,3	20,3	51,6	56,2	59,5	34,4	62,1	58,0	43,4	63,2	37,5	62,5	54,6	61,9	60,7	61,1	66,0	48,0
70. TAl26400	69,5	61,9	62,6	20,9	54,5	52,5	57,3	34,0	66,7	56,8	44,9	65,3	41,4	61,6	56,1	61,0	62,5	63,0	63,2	50,0

	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1. TC3698	51,4	53,4	51,7	31,6	48,5	38,8	52,7	25,4	54,5	53,0	53,0	32,7	36,2	53,3	56,0	31,1	26,5	32,4	26,5	56,9
2. AT1G03040	67,8	67,2	67,6	33,9	54,3	35,3	49,2	26,3	50,9	48,0	65,5	33,7	35,4	48,3	51,4	30,9	25,8	34,1	27,9	49,8
3. AT4G02590	72,3	73,6	73,3	32,8	57,1	37,7	50,3	27,5	52,6	49,5	67,5	34,3	35,4	50,8	52,7	32,8	27,8	33,6	28,7	51,2
4. AT2G24260	16,7	16,3	16,6	22,0	15,0	21,5	16,1	15,4	16,6	16,1	16,8	23,1	21,8	16,3	16,7	17,5	15,5	21,2	16,4	16,4
5. AT4G30980	36,1	35,9	36,2	40,8	34,8	44,8	36,2	26,7	36,5	36,1	35,4	41,8	44,3	35,6	39,8	33,9	25,5	36,2	28,6	37,6
6. AT5G58010	36,0	36,4	36,3	37,2	36,9	47,3	37,7	28,1	37,8	36,9	37,9	36,0	39,7	36,4	37,4	37,7	26,9	36,4	26,1	37,4
7. TC15501	60,7	59,0	58,0	36,2	47,9	37,1	54,0	30,8	53,8	53,8	58,7	38,2	37,6	53,0	52,5	32,7	29,4	35,0	30,7	52,0
8. TC19278	20,2	20,2	19,6	18,4	27,3	23,1	18,0	15,6	19,0	19,3	20,5	17,9	20,2	18,6	19,5	20,8	15,6	19,1	17,9	19,8
9. TC10015_part	67,6	69,0	68,9	31,7	55,3	38,5	51,3	25,5	53,4	50,0	64,1	32,8	35,7	50,6	51,8	32,9	26,0	32,9	26,7	52,8
10. DY268946	44,7	43,8	44,2	32,7	43,2	38,9	44,6	24,4	45,8	45,4	43,5	34,2	33,5	45,1	46,9	31,6	26,5	32,9	26,6	47,9
11. TA2544	36,1	35,7	36,5	41,8	29,6	41,3	35,3	30,2	35,4	33,8	35,1	45,9	42,6	35,0	35,1	39,3	31,1	42,5	32,1	35,3
12. TA3392	76,7	76,1	75,2	33,9	56,2	36,2	52,2	25,6	54,3	51,6	64,5	33,0	35,2	52,1	53,0	30,4	28,9	32,7	26,2	51,5
13. TAI2416	33,3	33,3	34,1	56,2	27,5	48,0	33,7	30,2	33,3	31,2	32,6	56,3	55,6	31,8	32,2	36,9	28,7	45,2	30,6	31,4
14. WO051050segID77	54,3	54,0	52,5	33,3	47,2	36,7	54,2	27,7	54,3	56,0	53,4	34,7	35,6	55,9	52,0	31,9	28,8	34,5	27,4	51,6
15. WO051050segID1671	70,7	70,3	68,9	32,0	53,2	34,4	47,6	27,4	49,4	48,5	62,0	34,8	35,0	48,7	48,7	32,1	28,7	33,1	29,3	48,6
16. TA55042	48,4	47,8	49,9	32,6	46,4	37,7	50,1	29,4	53,1	50,9	49,4	35,2	36,8	50,3	51,3	33,9	32,2	36,3	27,3	51,0
17. TC207545	64,4	65,1	65,6	30,9	77,6	37,8	52,6	23,9	55,6	52,0	61,8	32,0	32,9	53,0	54,1	28,6	23,9	31,3	24,9	53,2
18. TA13791	99,7	87,2	84,9	33,7	59,7	39,4	54,0	26,2	56,7	53,3	69,0	35,6	38,5	53,3	54,6	33,2	27,1	33,8	28,2	51,7
19. CO123623	50,5	49,5	49,2	30,9	52,2	39,3	51,1	25,4	53,2	50,3	47,7	31,5	33,6	52,0	50,6	31,1	29,0	30,1	27,3	50,9
20. DT543504	35,7	34,8	36,1	34,0	34,0	36,3	34,2	31,0	36,1	37,1	35,7	31,9	33,4	38,1	34,5	31,8	28,8	32,3	30,8	32,5
21. TC60118		86,8	84,6	33,7	59,4	39,1	53,7	26,2	56,4	53,0	68,7	35,6	38,5	53,0	54,3	33,2	27,5	33,6	28,2	51,4
22. TC60119	90,8		86,9	32,7	59,5	39,1	53,3	26,1	54,7	53,5	70,1	36,7	37,3	54,4	53,1	33,1	27,0	34,2	28,2	51,1
23. TC61833	90,5	92,5		33,5	59,3	37,7	52,6	26,5	55,9	53,3	71,1	36,4	36,8	53,9	51,1	34,5	25,4	34,7	27,4	52,0
24. TC67603	42,6	42,2	42,2		28,5	47,1	35,0	27,3	34,9	33,2	33,2	49,6	52,5	33,7	34,3	35,6	28,3	40,6	30,7	32,6
25. TC229602	65,0	64,5	63,9	35,8		38,2	47,7	22,5	50,5	50,7	58,2	29,0	32,3	50,0	50,8	28,2	21,3	28,4	22,7	49,3
26. TC205173	50,7	51,9	51,6	56,4	45,7		39,2	25,3	37,8	37,4	39,8	47,3	53,1	38,0	38,7	40,3	26,6	41,2	27,9	38,6
27. TA1140	65,7	66,0	65,4	45,2	54,2	52,2		27,1	84,2	66,0	51,2	35,8	37,5	67,0	52,5	33,3	27,2	31,8	28,2	52,1
28. TC140470	37,6	37,4	37,8	42,8	29,7	36,4	38,0		29,1	26,1	27,9	28,7	27,7	27,0	26,6	24,4	71,1	28,1	35,1	26,1

29. TA3490_part	68,3	66,4	66,6	44,7	57,2	49,6	87,3	37,4		68,0	53,2	34,4	36,5	68,5	54,0	32,7	28,3	33,3	27,6	53,7
30. CK293938	65,7	66,8	65,6	42,2	57,4	49,6	78,1	36,2	79,4		51,5	34,2	37,5	96,3	53,8	32,4	28,0	33,9	27,5	53,5
31. TC8633	80,5	81,6	82,3	44,3	63,2	51,3	67,3	38,5	68,9	67,5		36,0	34,9	52,5	53,8	33,0	29,2	33,6	26,9	52,6
32. TC7102	44,0	44,0	45,1	62,1	34,8	55,5	44,7	43,0	42,7	44,0	44,7		64,6	34,5	35,9	37,5	28,8	47,3	32,6	34,1
33. TC7103	48,8	49,1	48,3	65,8	40,1	62,5	48,8	41,0	47,5	47,0	46,5	72,9		38,2	37,8	41,1	28,9	46,2	29,6	38,3
34. TC12771	65,7	65,8	64,6	43,1	55,7	50,7	78,1	37,4	78,5	98,0	67,5	44,5	48,1		54,6	32,2	28,6	34,9	26,9	54,7
35. 0s03g58330	63,4	63,5	60,3	43,6	56,1	50,1	64,7	36,8	67,0	66,6	66,9	42,7	46,0	66,1		36,6	27,1	35,3	28,0	74,9
36. 0s06g08500	43,3	45,2	47,4	46,4	37,1	51,2	47,7	36,2	44,4	42,5	42,5	49,2	55,6	45,0	47,4		24,0	46,8	29,1	36,2
37. 0s06g09370	39,1	37,4	38,1	41,4	28,7	36,6	36,4	81,5	37,2	38,7	38,3	43,9	41,2	38,5	35,8	37,4		28,6	34,3	26,6
38. 0s02g55250	41,3	41,3	42,2	56,0	35,1	50,1	42,0	45,3	44,0	41,5	44,0	60,5	59:0	43,5	42,2	59,2	46,4		31,8	35,3
39. 0s02g35660	37,2	36,3	35,9	44,3	29,6	36,5	37,4	49,8	36,5	35,9	35,9	44,1	40,3	36,5	36,1	36,6	47,9	45,6		25,7
40. 0s07g08440	63,4	62,8	63,9	41,0	55,3	49,6	63,4	35,6	64,5	67,6	66,6	41,8	45,7	68,1	82,0	46,0	35,8	41,7	34,0	
41. 0s09g25040	38,1	38,1	38,7	54,5	31,1	47,9	38,7	42,5	36,3	36,3	37,5	58,5	54,9	37,1	37,3	47,3	43,7	56,3	44,8	37,5
42. Pt_scaff_II.416	89,1	88,2	85,6	42,4	66,0	50,4	65,0	37,0	67,3	67,0	82,1	44,5	48,3	66,7	63,3	44,7	39,5	43,3	36,6	64,3
43. Pt_scaff_70.65	63,1	63,7	63,7	43,3	52,3	50,7	63,7	38,7	63,4	65,6	63,1	44,5	47,0	65,9	60,7	44,7	42,9	44,2	37,8	60,7
44. Pt_scaff_XI11.403	64,1	62,3	63,2	44,0	50,6	54,0	62,9	41,0	65,0	63,5	63,5	44,9	47,5	63,2	59,9	43,6	41,8	44,9	38,2	59,0
45. Pt_scaff28.86	39,3	39,9	39,5	47,8	31,8	42,1	39,9	47,6	38,6	40,6	41,4	47,6	45,8	40,1	36,4	41,2	49,6	47,6	45,6	35,3
46. Pt_TC63334	88,8	88,2	86,2	42,4	66,7	51,0	63,1	37,4	67,0	64,3	82,8	45,1	48,8	64,7	64,0	44,1	38,3	43,5	35,3	65,0
47. TA14134	50,2	48,9	46,5	43,6	42,8	54,9	50,2	35,1	49,2	48,9	51,7	43,6	45,5	49,8	47,7	46,6	36,8	44,7	36,8	46,8
48. TA5414	42,7	41,0	41,7	67,1	33,5	55,1	42,1	48,6	41,5	42,1	41,2	69,4	63,9	40,8	38,5	47,2	44,8	57,1	45,4	39,3
49. TA3263	89,1	88,5	86,6	43,1	66,7	51,0	63,1	37,4	67,3	64,7	82,8	45,6	49,1	65,0	64,0	44,4	38,7	43,8	35,1	64,3
50. TA2825	90,1	90,1	89,2	43,1	67,2	51,0	66,0	35,8	67,2	66,2	83,4	46,0	49,4	66,2	63,2	45,0	37,9	43,3	36,6	65,2
51. TA1616	66,3	65,0	65,0	42,9	55,2	51,0	65,6	38,3	63,2	65,6	66,0	47,2	47,8	66,6	63,2	48,5	40,8	46,3	36,3	66,0
52. TA21665	65,7	66,4	66,6	42,9	54,4	51,3	73,5	36,6	76,4	84,8	64,2	43,8	48,6	84,9	65,9	44,7	35,6	43,3	37,2	65,9
53. TC172581	66,8	67,4	65,6	42,2	55,6	50,7	76,1	38,5	76,0	92,1	66,1	43,8	48,3	92,8	65,8	44,4	38,5	41,5	36,6	64,5
54. AK247217	82,5	84,2	84,3	43,3	63,9	49,6	67,6	38,0	67,9	69,2	95,0	42,7	47,3	67,5	64,6	43,6	37,9	44,2	35,9	64,6
55. TC104646	62,7	61,5	61,6	43,1	59,2	49,3	62,1	35,1	66,6	67,6	62,9	40,2	44,2	67,1	77,6	45,2	35,8	40,1	34,0	85,0
56. DV982110	57,8	57,9	57,7	44,5	56,2	58,2	58,5	35,3	57,9	55,7	55,6	43,6	50,4	54,4	57,5	46,0	35,6	43,5	34,9	59,0
57. WO051050segID516	43,0	43,6	43,6	50,5	35,0	46,0	41,9	45,9	41,3	41,9	43,2	52,9	48,8	42,1	42,5	44,3	44,6	51,6	42,6	41,9
58. TC68930	50,5	48,0	49,5	45,2	43,5	57,9	50,2	37,0	48,9	47,1	47,4	45,1	50,4	47,1	49,8	47,7	37,0	45,8	35,9	46,8
59. TA30646	83,2	84,2	84,3	42,9	63,9	49,6	68,3	37,6	67,5	68,5	95,4	42,2	46,8	68,5	62,3	43,1	39,3	45,6	36,3	65,9
60. TA28621	67,8	68,4	66,6	42,4	54,9	51,3	77,1	38,7	76,6	93,8	66,8	43,6	48,6	94,4	66,8	45,0	38,5	41,5	36,5	64,1
61. TA34455	42,3	43,4	41,8	63,1	34,2	53,0	44,5	45,3	42,3	44,3	42,7	83,2	71,4	43,6	44,3	47,4	44,4	60,0	43,1	41,6
62. TA37666	48,7	48,7	47,4	63,7	39,8	61,0	49,5	42,4	48,2	49,5	49,5	71,1	88,9	50,3	48,2	54,7	41,8	55,8	38,5	47,6
63. TC253044	52,8	54,9	50,8	40,3	47,6	48,4	50,3	56,5	55,9	54,7	54,0	41,3	45,7	55,7	52,7	42,5	50,0	40,6	38,4	52,9
64. TabHLH11	62,4	62,2	61,6	40,5	58,7	50,4	63,1	35,8	66,2	66,9	65,6	41,8	45,2	67,8	86,4	43,3	36,8	41,0	36,3	80,5
65. TA46156	85,1	85,2	83,3	44,7	67,8	53,4	67,3	38,0	70,0	70,9	81,1	43,3	48,3	70,8	66,0	44,1	37,0	42,4	35,7	66,2
66. GSVIVT00017237001	45,9	45,0	45,9	64,9	35,9	61,2	44,0	42,6	42,1	42,8	46,2	68,2	65,1	43,8	43,3	52,4	43,1	58,5	42,4	44,7
67. TA46194	73,2	72,5	72,2	44,3	57,8	54,3	70,6	38,7	72,8	74,8	73,2	44,9	49,6	75,1	68,4	45,8	39,1	45,1	37,6	68,7
68. GSVIVT00016367001	36,9	36,5	38,1	44,3	29,9	41,6	39,0	50,3	36,9	35,3	38,8	45,4	44,3	35,5	36,5	39,6	50,5	46,2	49,4	35,9
69. TA139285	60,7	61,5	61,0	40,7	56,1	49,0	64,7	35,6	66,2	68,9	64,2	41,1	46,8	68,1	90,5	45,5	38,1	40,4	34,0	84,0
70. TAl26400	62,7	61,8	63,6	42,6	57,9	49,3	62,4	34,9	65,9	66,9	62,3	40,0	45,2	66,1	76,2	44,1	35,6	40,8	35,3	82,9

	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
1. TC3698	30,0	50,8	48,7	47,8	28,7	50,8	31,8	29,7	51,4	54,0	50,6	52,3	50,2	54,4	58,0	45,1	33,3	33,4	54,7	50,8
2. AT1G03040	29,5	65,7	48,6	47,5	29,9	66,7	31,1	31,1	66,7	70,0	52,4	46,8	49,2	66,5	50,6	40,2	33,0	31,8	67,0	49,2
3. AT4G02590	31,7	71,8	52,1	49,3	30,9	72,8	31,6	33,6	73,1	75,2	54,0	49,4	53,2	70,2	53,4	43,1	33,3	32,5	69,9	52,7
4. AT2G24260	21,4	16,7	16,6	15,5	18,1	17,0	18,9	23,0	17,0	16,4	17,3	15,6	15,4	17,0	16,6	18,5	20,8	19,9	16,5	15,7
5. AT4G30980	33,8	35,5	36,1	33,6	29,6	36,4	34,9	41,6	36,8	37,5	36,4	36,1	36,9	36,4	41,0	42,2	34,9	36,4	36,6	35,7
6. AT5G58010	31,9	34,6	34,2	34,1	31,3	35,8	36,3	35,6	35,8	35,1	33,9	36,2	35,1	35,4	38,3	46,5	33,5	39,9	36,2	34,7
7. TC15501	33,5	59,7	55,6	55,5	30,7	62,0	30,8	35,2	61,5	61,2	59,3	51,0	51,4	61,1	51,6	41,5	36,4	30,0	60,3	51,7
8. TC19278	17,0	19,7	17,0	18,0	22,7	19,7	29,4	19,0	20,5	20,3	18,3	18,8	18,4	19,8	21,0	25,1	20,2	30,4	20,1	18,4
9. TC10015_part	30,6	67,5	50,4	48,7	27,4	68,5	30,9	31,0	68,8	70,7	53,3	48,7	50,0	65,9	51,8	42,4	32,8	31,8	66,2	50,0
10. DY268946	31,5	42,9	54,2	51,6	29,6	44,0	36,8	33,1	44,0	43,7	57,3	42,4	43,9	41,7	47,0	40,7	35,5	35,9	44,5	44,9
11. TA2544	43,3	36,2	36,9	36,4	33,1	36,4	32,5	45,2	36,6	36,1	37,0	33,5	35,9	35,7	35,1	37,6	40,8	34,2	34,5	36,1
12. TA3392	30,8	76,3	49,3	48,6	28,3	78,0	30,2	31,0	78,0	76,7	52,6	53,2	52,5	68,3	51,8	42,8	33,4	31,3	68,0	53,2
13. TAI2416	47,5	32,7	32,7	33,3	32,5	33,5	31,2	98,3	32,8	33,5	33,5	30,3	32,2	32,6	32,9	35,2	37,7	31,7	32,2	32,2
14. WO051050segID77	32,3	54,0	63,0	61,3	29,3	54,6	32,1	32,0	54,6	54,3	63,7	54,0	54,0	55,3	52,0	41,4	34,8	31,2	55,0	53,5
15. WO051050segID1671	33,4	70,4	50,5	50,7	30,9	70,3	28,9	32,9	70,6	73,4	51,6	46,5	48,3	65,7	48,6	38,5	33,5	29,4	66,0	48,6
16. TA55042	31,1	47,8	61,3	60,1	30,4	48,1	30,9	33,3	48,1	50,0	62,5	47,4	50,3	49,2	51,5	40,4	31,6	33,0	49,5	50,9
17. TC207545	28,8	67,3	48,3	44,8	27,2	68,3	32,8	30,2	68,7	69,8	48,5	50,7	51,6	64,1	56,7	49,3	31,8	35,2	64,4	51,3
18. TA13791	32,5	80,7	50,1	50,7	30,7	82,0	32,6	33,3	82,7	82,1	54,9	53,4	54,2	71,8	54,9	43,3	35,4	32,8	72,4	54,2
19. CO123623	29,6	50,3	58,4	56,5	26,2	51,5	32,1	33,1	51,5	48,1	61,3	47,6	48,1	45,8	53,4	47,8	32,2	33,1	46,1	49,7
20. DT543504	29,0	33,3	33,3	32,7	40,8	32,9	36,1	30,8	33,5	35,1	33,1	35,0	36,4	34,1	34,6	41,2	31,0	37,2	35,3	37,2
21. TC60118	32,3	80,4	49,9	50,4	30,7	81,7	32,6	33,3	82,4	81,8	54,7	53,1	53,9	71,5	54,6	43,0	35,2	32,3	72,1	53,9
22. TC60119	31,4	80,5	49,9	48,6	30,1	81,8	31,2	33,6	82,4	82,8	53,3	53,8	54,1	73,5	51,7	42,9	35,1	32,4	73,8	55,0
23. TC61833	31,2	78,9	49,1	48,4	30,1	80,5	30,2	34,7	81,2	81,6	53,3	53,1	55,0	74,3	52,3	43,2	35,2	33,3	74,3	54,9
24. TC67603	43,3	33,3	33,2	32,0	33,3	34,0	31,5	56,8	33,9	32,8	34,2	32,5	33,0	32,7	33,9	37,6	36,3	33,2	32,5	33,6

25. TC229602	26,2	61,0	43,8	41,6	25,1	62,3	32,9	27,7	62,7	64,2	47,9	47,0	49,0	59,8	52,5	48,2	29,4	34,5	60,1	49,0
26. TC205173	41,4	38,3	36,5	36,4	30,9	38,3	36,8	47,9	38,3	39,6	36,5	37,0	38,1	38,2	38,2	48,5	36,9	40,7	38,0	37,3
27. TA1140	31,2	54,7	54,0	50,8	27,7	52,7	31,9	34,3	52,7	52,9	53,7	61,7	64,3	55,4	52,8	41,1	31,4	32,8	54,0	64,9
28. TC140470	28,2	26,2	28,1	28,7	31,9	25,2	24,0	31,2	25,3	24,9	27,4	27,6	27,4	29,0	26,1	27,3	27,7	25,0	28,6	27,4
29. TA3490_part	30,3	54,9	53,0	52,6	29,0	54,3	31,7	33,3	54,9	55,0	53,7	66,0	67,0	55,5	56,8	45,3	32,3	33,6	55,5	67,3
30. CK293938	30,0	53,0	52,7	50,9	28,6	52,0	30,9	31,2	52,6	54,0	54,1	76,4	87,9	53,3	56,8	41,2	32,7	32,7	53,5	88,9
31. TC8633	31,4	69,9	52,7	50,7	30,5	71,2	34,0	32,1	71,5	72,4	52,6	51,5	51,9	90,8	52,7	43,9	35,0	31,9	91,1	51,9
32. TC7102	48,4	35,3	33,5	34,4	32,9	35,7	32,8	56,1	35,7	36,7	36,8	33,2	35,6	35,7	34,3	36,2	38,0	33,4	35,2	35,3
33. TC7103	44,7	34,3	34,8	36,7	32,9	35,7	34,1	55,3	36,2	37,2	37,7	36,6	37,8	36,2	37,0	41,0	37,6	37,0	36,0	38,5
34. TC12771	30,0	53,6	53,0	50,6	28,8	51,7	31,4	31,9	52,3	54,0	54,7	77,7	89,2	54,1	57,4	40,9	32,1	33,1	53,8	90,5
35. 0s03g58330	32,3	54,3	50,1	49,0	27,4	51,2	33,5	31,8	51,2	55,0	52,8	51,1	52,1	53,6	72,9	44,2	34,1	34,6	54,0	52,7
36. 0s06g08500	38,0	32,9	32,1	31,3	27,7	34,3	32,3	37,9	34,6	33,9	34,5	29,6	33,1	33,1	33,3	36,6	32,8	32,0	32,8	32,7
37. 0s06g09370	28,6	27,6	31,2	30,4	32,1	27,9	23,9	27,8	28,1	26,7	30,1	26,3	28,6	28,5	27,4	28,2	27,5	25,4	28,4	27,6
38. 0s02g55250	46,4	33,4	34,2	34,4	32,8	33,7	33,5	45,8	34,7	34,1	35,8	33,3	34,2	33,5	34,0	36,1	36,8	34,3	34,6	34,0
39. 0s02g35660	30,4	26,4	29,1	27,2	31,9	26,1	25,7	30,5	26,4	28,2	27,7	25,5	27,8	27,3	26,5	27,9	29,1	26,6	27,6	28,1
40. 0s07908440	31,4	53,4	50,0	47,8	27,7	52,3	31,7	31,7	52,3	56,2	55,8	51,9	50,8	51,4	81,3	45,6	33,6	31,9	52,6	50,8
41.Os09g25040		30,6	31,2	31,2	31,1	30,9	29,1	47,7	31,7	31,5	31,6	29,4	30,5	31,7	31,1	32,5	37,2	31,3	31,5	30,4
42. Pt_scaff_II.416	38,3		50,3	49,6	29,1	93,7	30,5	33,0	93,0	88,9	52,1	52,5	54,9	72,7	53,4	44,4	36,5	32,4	73,0	54,5
43. Pt_scaff_70.65	38,1	63,1		88,4	29,5	50,0	31,0	32,4	49,7	51,6	73,2	51,5	51,7	52,3	49,4	42,3	34,6	31,6	52,0	52,0
44. Pt_scaff_XIII.403	38,5	63,2	90,7		27,9	49,6	28,2	32,3	49,3	49,6	71,3	49,9	50,4	50,6	48,8	40,0	34,8	30,5	50,6	51,3
45. Pt_scaff28.86	46,1	39,3	40,4	40,6		29,8	31,8	33,2	30,3	28,9	29,8	28,7	27,9	30,1	27,6	31,1	33,6	32,2	30,1	28,5
46. Pt_TC63334	38,1	95,3	63,4	61,4	40,4		30,8	33,6	98,7	89,2	52,3	52,5	53,9	74,0	53,7	44,6	35,2	33,1	74,3	53,9
47. TA14134	38,1	48,3	49,2	47,6	42,1	49,2		30,1	30,8	31,4	31,4	29,6	30,5	33,2	31,4	39,1	33,7	49,4	33,4	30,0
48. TA5414	60,5	42,1	41,0	42,1	47,9	42,7	39,7		32,7	33,8	32,4	30,3	32,5	32,4	31,9	35,3	38,1	30,2	31,8	32,5
49. TA3263	38,7	95,0	63,4	61,4	41,0	99,0	49,8	42,1		89,2	52,6	53,1	54,5	74,3	54,3	44,9	35,8	33,6	74,6	54,5
50. TA2825	39,1	93,3	64,7	62,9	39,0	93,0	50,2	41,2	93,3		53,6	53,6	54,9	76,3	54,0	44,8	35,0	32,0	76,3	54,9
51. TA1616	39,5	65,6	83,7	80,5	41,2	64,1	50,3	40,6	64,1	66,3		51,1	52,4	54,1	54,8	42,5	35,7	32,1	54,1	53,6
52. TA21665	37,7	67,0	63,7	63,2	37,7	66,7	48,0	41,2	67,0	66,6	65,0		75,2	52,3	53,8	43,0	32,0	32,5	52,0	76,5
53. TC172581	38,1	68,4	64,7	63,5	39,9	66,8	49,8	42,1	67,1	68,4	65,6	82,2		53,3	54,0	43,5	34,0	31,8	52,1	96,7
54. AK247217	37,9	83,8	65,3	63,8	42,8	84,4	50,8	41,9	84,4	86,1	66,9	65,6	66,8		52,8	44,3	35,0	32,8	98,3	53,8
55. TC104646	37,1	63,3	57,4	57,8	36,0	63,7	44,6	38,9	64,0	62,9	61,7	65,9	65,8	62,9		48,5	34,2	34,4	53,5	55,0
56. DV982110	38,9	58,3	54,1	50,3	40,1	58,3	51,1	41,7	58,7	57,5	53,7	57,4	53,3	55,0	60,3		54,2	40,7	45,0	43,2
57. WO051050segID516	52,3	44,7	44,5	44,7	49,5	42,5	42,8	51,7	43,0	43,0	46,0	41,7	43,0	42,3	42,3	55,7		37,2	34,6	33,7
58. TC68930	40,9	49,8	48,0	46,1	45,0	50,8	63,7	40,4	51,1	48,9	50,2	48,3	50,2	49,5	47,7	53,2	47,9		32,9	31,7

59. TA30646	37,9	84,1	65,3	63,8	42,8	84,8	49,8	41,7	84,8	85,8	66,9	65,2	66,8	99,0	63,6	56,6	42,1	49,2		53,0
60. TA28621	37,7	68,8	65,3	63,8	40,4	67,4	48,3	41,2	67,8	69,1	66,0	83,6	97,7	67,1	65,5	53,6	43,8	48,3	67,4	
61. TA34455	59,5	42,5	44,7	46,5	46,1	41,2	42,7	66,9	40,9	42,5	43,0	40,9	43,2	41,4	40,0	42,5	53,8	44,5	41,2	42,7
62. TA37666	52,3	49,7	51,0	49,2	43,9	50,0	47,1	63,5	50,3	47,6	48,2	49,7	48,4	47,6	44,5	46,1	48,8	51,0	48,2	49,5
63. TC253044	36,5	54,0	49,8	50,3	41,4	53,3	46,8	39,1	54,0	54,5	49,1	54,4	53,6	51,7	50,0	55,2	40,0	51,7	51,0	53,9
64. TabHLH11	36,5	62,3	58,3	59,0	36,4	63,0	45,5	38,2	63,3	63,2	59,5	65,9	66,1	65,6	83,3	60,5	43,0	46,2	65,2	64,5
65. TA46156	37,5	85,0	64,7	61,4	39,0	85,3	50,5	40,2	85,7	87,0	65,0	66,9	67,8	82,5	65,1	60,2	42,5	48,9	82,5	69,4
66. GSVIVT00017237001	56,5	44,3	45,9	48,1	47,6	45,0	44,0	69,2	45,2	44,7	45,0	43,1	44,7	44,7	41,4	45,7	55,9	47,4	44,5	45,2
67. TA46194	39,7	73,8	70,4	69,2	40,1	73,5	49,5	42,1	73,8	73,5	74,8	72,5	73,2	74,4	66,8	56,9	44,3	50,5	74,4	73,5
68. GSVIVT00016367001	45,9	38,8	40,6	41,2	63,9	36,7	39,0	45,6	36,3	37,9	37,9	36,7	37,3	38,8	34,6	37,3	44,5	40,6	38,6	37,1
69. TA139285	37,7	61,7	58,3	58,4	38,4	61,7	47,1	38,2	61,7	61,2	64,1	66,2	65,5	64,2	82,5	59,6	42,8	48,6	64,6	65,8
70. TAI26400	36,7	64,0	58,0	55,7	36,4	64,7	45,8	41,5	65,0	64,9	62,0	66,2	64,5	61,9	92,3	59,3	42,5	48,0	61,9	65,5

	61	**62**	**63**	**64**	**65**	**66**	**67**	**68**	**69**	**70**
1. TC3698	31,4	35,3	37,5	56,1	53,5	34,9	54,6	25,8	58,1	55,7
2. AT1G03040	32,8	36,4	37,3	50,6	63,2	33,5	53,4	28,0	50,8	49,4
3. AT4G02590	33,4	38,9	37,0	53,7	69,0	34,8	56,3	28,5	53,7	50,7
4. AT2G24260	20,6	21,8	14,8	17,0	16,3	22,1	16,5	16,3	16,8	16,5
5. AT4G30980	41,3	43,2	35,3	39,5	38,6	39,7	37,3	29,5	40,2	40,5
6. AT5G58010	36,6	40,5	36,6	37,3	36,1	36,2	35,9	28,6	38,4	37,5
7. TC15501	37,0	37,5	34,0	51,1	59,6	34,8	64,3	30,7	53,4	49,7
8. TC19278	17,0	20,6	24,1	19,9	21,3	20,0	20,1	19,5	20,2	20,1
9. TC10015_part	32,7	35,1	36,8	53,9	68,5	33,2	54,4	26,9	51,8	51,6
10. DY268946	34,0	34,6	33,1	49,0	45,3	36,4	50,0	30,2	46,1	45,5
11. TA2544	46,6	45,3	31,7	32,9	36,3	42,9	37,1	33,9	36,2	35,0
12. TA3392	34,1	36,5	37,5	53,7	73,3	34,2	57,6	28,0	54,4	52,6
13. TAI2416	53,7	52,9	28,8	32,9	33,3	58,2	33,3	31,9	31,8	33,3
14. WO051050segID77	35,8	33,9	36,7	49,2	56,2	33,4	60,3	26,9	52,6	51,2
15. WO051050segID1671	33,8	34,8	33,8	48,9	66,8	34,0	57,0	29,8	49,6	48,9
16. TA55042	35,3	36,8	35,1	49,0	51,4	34,2	56,0	28,4	52,0	51,2
17. TC207545	29,6	32,8	39,7	54,5	67,5	33,2	55,4	26,1	54,6	55,6
18. TA13791	36,0	38,0	36,3	54,7	75,6	36,5	59,8	29,1	54,6	53,7
19. CO123623	29,8	33,3	36,7	53,0	49,4	30,9	55,6	27,3	52,5	51,9
20. DT543504	32,0	33,9	41,4	37,9	36,0	31,4	34,5	37,8	34,8	35,2

21. TC60118	36,0	38,0	36,3	54,4	75,2	36,5	59,5	29,1	54,2	53,4
22. TC60119	34,8	37,5	36,4	53,9	75,7	35,6	59,4	29,0	53,9	51,2
23. TC61833	34,0	36,7	34,1	53,1	73,4	35,2	59,2	28,6	53,8	52,5
24. TC67603	46,9	51,7	30,2	32,6	34,9	51,4	34,0	31,0	32,4	33,9
25. TC229602	27,7	31,3	34,7	51,4	63,3	29,4	53,2	23,3	51,4	50,0
26. TC205173	42,1	48,7	35,1	38,1	39,7	52,3	40,4	30,7	37,6	38,8
27. TA1140	34,5	37,4	36,7	52,1	54,3	35,5	61,0	27,7	54,6	52,8
28. TC140470	30,3	28,1	55,1	25,1	24,9	29,7	28,1	34,4	25,2	25,5
29. TA3490_part	33,6	36,7	36,5	54,6	58,5	34,0	61,3	26,3	55,1	54,5
30. CK293938	34,3	37,7	36,6	52,6	55,9	34,3	65,0	27,8	55,8	54,2
31. TC8633	33,7	36,8	35,2	53,7	71,7	36,0	59,3	29,4	54,1	50,9
32. TC7102	71,8	63,0	31,4	36,8	34,2	54,6	35,6	30,0	34,8	34,0
33. TC7103	62,4	83,3	33,6	39,1	36,0	53,1	38,0	29,6	38,5	36,9
34. TC12771	34,6	38,1	36,6	53,7	57,2	33,9	66,2	28,7	55,8	54,2
35. 0s03g58330	36,6	37,6	34,9	80,3	53,5	34,5	58,3	27,8	87,2	69,4
36. 0s06g08500	39,0	40,0	30,1	34,1	32,6	40,3	33,5	27,5	35,9	32,3
37. 0s06g09370	30,3	28,4	45,0	26,8	27,4	29,2	27,7	34,1	28,0	25,9
38. 0s02g55250	46,1	44,2	31,3	33,8	33,6	46,3	34,4	30,3	34,6	34,8
39. 0s02g35660	31,3	29,8	29,5	27,1	25,9	31,3	27,7	33,7	27,2	27,6
40. 0s07g08440	33,6	36,3	35,3	74,9	53,0	35,8	57,1	27,0	79,7	76,2
41. 0s09g25040	45,6	43,6	27,9	31,1	30,5	47,3	31,6	30,4	32,3	31,4
42. Pt_scaff_II.416	34,1	35,6	35,5	53,9	74,1	34,4	61,5	28,9	55,3	52,2
43. Pt_scaff_70.65	34,1	37,6	35,0	48,8	53,6	35,0	58,9	28,0	49,5	50,1
44. Pt_scaff_XIII.403	34,7	36,5	34,2	47,5	50,3	35,8	57,4	29,5	47,6	48,1
45. Pt_scaff28.86	32,2	32,3	31,2	28,1	28,3	34,3	31,0	49,3	29,8	26,4
46. Pt_TC63334	33,9	37,2	34,8	54,8	76,6	35,1	60,0	28,5	55,6	53,1
47. TA14134	30,4	35,1	32,5	32,0	32,4	33,7	32,5	29,4	32,9	32,2
48. TA5414	52,7	53,3	29,2	32,9	33,8	57,9	33,8	32,1	32,6	32,8
49. TA3263	33,8	37,7	34,8	55,1	76,9	35,6	60,0	28,3	55,6	53,4
50. TA2825	34,6	35,8	35,6	55,8	77,7	35,8	62,0	28,0	55,5	53,7
51. TA1616	35,0	36,8	34,7	51,4	56,6	34,4	62,5	29,2	55,0	54,4
52. TA21665	32,3	36,3	35,6	52,1	52,9	33,2	60,5	27,6	53,6	52,7
53. TC172581	34,2	37,7	37,1	51,9	57,3	35,3	64,2	28,6	53,1	51,6
54. AK247217	34,2	36,4	34,4	54,0	73,9	35,3	60,2	30,1	53,3	50,9

55. TC104646	34,4	36,3	35,7	74,4	54,6	34,2	58,1	26,3	77,6	91,2
56. DV982110	34,7	38,2	41,9	45,8	45,4	37,3	43,8	29,7	46,7	45,9
57. WO051050segID516	38,0	37,3	29,6	33,4	35,4	40,4	36,1	29,6	34,8	33,3
58. TC68930	31,5	36,6	35,7	32,2	32,1	35,4	32,3	30,7	34,3	34,6
59. TA30646	32,0	36,7	34,4	54,7	73,9	35,5	60,2	29,5	54,9	51,2
60. TA28621	33,6	39,1	37,4	51,4	56,2	35,3	63,5	28,7	53,4	52,2
61. TA34455		61,4	30,3	35,9	33,5	50,6	35,4	32,2	34,9	34,6
62. TA37666	70,7		32,2	39,2	36,2	52,3	38,4	30,4	37,4	37,9
63. TC253044	40,9	47,1		34,6	36,5	33,0	37,2	30,6	35,8	35,8
64. TabHLH11	41,6	45,5	51,4		52,5	33,3	55,7	27,7	82,7	70,7
65. TA46156	41,6	47,6	53,8	66,4		35,1	65,7	27,0	55,3	52,9
66. GSVIVT00017237001	67,1	64,4	43,5	41,9	45,9		36,0	31,6	32,5	35,6
67. TA46194	45,2	50,5	51,1	66,1	75,7	45,0		29,3	57,5	56,5
68. GSVIVT00016367001	46,8	42,7	41,4	35,1	36,1	44,5	37,5		26,9	26,1
69. TA139285	43,2	44,0	54,8	89,5	66,4	40,0	68,1	37,3		72,8
70. TAI26400	40,7	46,1	53,8	79,6	65,1	41,9	66,5	35,7	80,4	

Tabela E3: Resultados MatGAT para a similaridade e identidade global nos domínios HLH das sequências de polipeptídeos semelhante à bHLH11.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. AT2G24260		98,0	92,2	98,0	98,0	98,0	98,0	98,0	98,0	94,1	94,1	96,1	86,3	98,0	88,9	96,1	98,0	98,0
2. AT4G30980	100,0		94,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	94,1	96,1	96,1
3. AT5G58010	98,0	98,0		92,2	92,2	92,2	92,2	92,2	92,2	96,1	96,1	90,2	84,3	92,2	83,3	90,2	92,2	92,2
4. TAI2416	98,0	98,0	98,0		100,0	100,0	100,0	100,0	100,0	96,1	96,1	98,0	88,2	100,0	90,7	98,0	98,0	98,0
5. TA5414	98,0	98,0	98,0	100,0		100,0	100,0	100,0	100,0	96,1	96,1	98,0	88,2	100,0	90,7	98,0	98,0	98,0
6. TC205173	98,0	98,0	98,0	100,0	100,0		100,0	100,0	100,0	96,1	96,1	98,0	88,2	100,0	90,7	98,0	98,0	98,0
7. GSVIVT00017237001	98,0	98,0	98,0	100,0	100,0	100,0		100,0	100,0	96,1	96,1	98,0	88,2	100,0	90,7	98,0	98,0	98,0
8. TC7102	98,0	98,0	98,0	100,0	100,0	100,0	100,0		100,0	96,1	96,1	98,0	88,2	100,0	90,7	98,0	98,0	98,0
9. TA34455	98,0	98,0	98,0	100,0	100,0	100,0	100,0	100,0		96,1	96,1	98,0	88,2	100,0	90,7	98,0	98,0	98,0
10. TC7103	98,0	98,0	98,0	100,0	100,0	100,0	100,0	100,0	100,0		100,0	94,1	88,2	96,1	87,0	94,1	94,1	94,1
11. TA37666	98,0	98,0	98,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0		94,1	88,2	96,1	87,0	94,1	94,1	94,1
12. TC67603	96,1	96,1	96,1	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0		86,3	98,0	88,9	96,1	96,1	96,1
13. 0s06g08500	90,2	90,2	90,2	92,2	92,2	92,2	92,2	92,2	92,2	92,2	92,2	90,2		88,2	79,6	86,3	86,3	86,3
14. 0s02g55250	98,0	98,0	98,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	98,0	92,2		90,7	98,0	98,0	98,0
15. 0s09g25040	88,9	88,9	88,9	90,7	90,7	90,7	90,7	90,7	90,7	90,7	90,7	88,9	83,3	90,7		88,9	88,9	88,9

16. TA2544	98,0	98,0	98,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	98,0	92,2	100,0	90,7		96,1	96,1	
17. DV982110	98,0	98,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	98,0	92,2	100,0	90,7	100,0		100,0	
18. WO051050segID516	98,0	98,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	98,0	92,2	100,0	90,7	100,0	100,0		
19. AT4G02590	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	96,1	96,1	
20. TC10015_part	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	96,1	96,1	
21. AT1 G03040	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	96,1	96,1	
22. TA13791	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
23. TC60118	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
24. TC60119	94,1	94,1	94,1	94,1	94,1	94,1	94,1	94,1	94,1	94,1	94,1	92,2	86,3	94,1	85,2	94,1	94,1	94,1	
25. TC61833	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	96,1	96,1	
26. TA3392	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
27. WO051050segID1671	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	96,1	96,1	
28. Pt_TC63334	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	96,1	96,1	
29. TA3263	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
30. Pt_scaff_II.416	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
31. TA2825	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
32. TC207545	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
33. TC229602	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
34. TA46156	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
35. AK247217	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	96,1	96,1	
36. TA30646	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	96,1	96,1	
37. TC8633	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	96,1	96,1	
38. TC15501	98,0	98,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	98,0	92,2	100,0	90,7	100,0	100,0	100,0	
39. TA1140	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
40. TA3490_part	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
41. CK293938	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
42. TC12771	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
43. TC172581	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
44. TA28621	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
45. TA21665	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
46. TA46194	98,0	98,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	98,0	92,2	100,0	90,7	100,0	100,0	100,0	
47. DY268946	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
48. CO123623	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0	
49. Pt_scaff_70.65	92,2	92,2	92,2	92,2	92,2	92,2	92,2	92,2	92,2	92,2	92,2	94,1	84,3	92,2	83,3	92,2	92,2	92,2	

50. Pt_scaff_XIII.403	92,2	92,2	92,2	92,2	92,2	92,2	92,2	92,2	92,2	92,2	92,2	94,1	84,3	92,2	83,3	92,2	92,2	92,2
51. TA1616	96,1	96,1	96,1	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0
52. TA55042	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	98,0	98,0
53. WO051050segID77	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	96,1	96,1
54. TC3698	96,1	96,1	98,0	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	98,0	98,0
55. 0s03g58330	98,0	98,0	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	100,0	100,0
56. TA139285	98,0	98,0	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	100,0	100,0
57. TabHLH11	96,1	96,1	98,0	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	98,0	98,0
58. TC104646	98,0	98,0	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	100,0	100,0
59. TAI26400	98,0	98,0	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	100,0	100,0
60. 0s07g08440	98,0	98,0	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	100,0	100,0
61. 0s02g35660	96,1	96,1	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	96,1	90,2	98,0	88,9	98,0	100,0	100,0
62. DT543504	94,1	96,1	98,0	96,1	96,1	96,1	96,1	96,1	96,1	98,0	98,0	94,1	90,2	96,1	87,0	96,1	96,1	96,1
63. Pt_scaff28.86	94,1	96,1	98,0	96,1	96,1	96,1	96,1	96,1	96,1	98,0	98,0	94,1	90,2	96,1	88,9	96,1	96,1	96,1
64. GSVIVT00016367001	96,1	98,0	100,0	98,0	98,0	98,0	98,0	98,0	98,0	100,0	100,0	96,1	92,2	98,0	88,9	98,0	98,0	98,0
65. TC19278	68,6	70,6	72,5	70,6	70,6	70,6	70,6	70,6	70,6	72,5	72,5	70,6	72,5	70,6	63,0	70,6	70,6	70,6
66. TA14134	94,1	94,1	94,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	88,2	96,1	87,0	96,1	96,1	96,1
67. TC68930	98,0	98,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	98,0	92,2	100,0	90,7	100,0	100,0	100,0
68. TC140470	94,1	96,1	98,0	96,1	96,1	96,1	96,1	96,1	96,1	98,0	98,0	94,1	90,2	96,1	87,0	96,1	96,1	96,1
69. TC253044	94,1	96,1	98,0	96,1	96,1	96,1	96,1	96,1	96,1	98,0	98,0	94,1	90,2	96,1	87,0	96,1	96,1	96,1
70. 0s06g09370	92,2	94,1	96,1	94,1	94,1	94,1	94,1	94,1	94,1	96,1	96,1	92,2	88,2	94,1	85,2	94,1	94,1	94,1

	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1. AT2G24260	82,4	82,4	80,4	84,3	84,3	82,4	84,3	84,3	80,4	82,4	84,3	82,4	84,3	84,3	84,3	84,3	82,4	84,3
2. AT4G30980	80,4	80,4	82,4	82,4	82,4	80,4	82,4	82,4	80,4	80,4	82,4	80,4	82,4	82,4	82,4	82,4	80,4	82,4
3. AT5G58010	76,5	76,5	78,4	78,4	78,4	76,5	78,4	78,4	76,5	76,5	78,4	76,5	78,4	78,4	78,4	78,4	76,5	78,4
4. TAI2416	82,4	82,4	80,4	84,3	84,3	82,4	84,3	84,3	80,4	82,4	84,3	82,4	84,3	84,3	84,3	84,3	82,4	84,3
5. TA5414	82,4	82,4	80,4	84,3	84,3	82,4	84,3	84,3	80,4	82,4	84,3	82,4	84,3	84,3	84,3	84,3	82,4	84,3
6. TC205173	82,4	82,4	80,4	84,3	84,3	82,4	84,3	84,3	80,4	82,4	84,3	82,4	84,3	84,3	84,3	84,3	82,4	84,3
7. GSVIVT00017237001	82,4	82,4	80,4	84,3	84,3	82,4	84,3	84,3	80,4	82,4	84,3	82,4	84,3	84,3	84,3	84,3	82,4	84,3
8. TC7102	82,4	82,4	80,4	84,3	84,3	82,4	84,3	84,3	80,4	82,4	84,3	82,4	84,3	84,3	84,3	84,3	82,4	84,3
9. TA34455	82,4	82,4	80,4	84,3	84,3	82,4	84,3	84,3	80,4	82,4	84,3	82,4	84,3	84,3	84,3	84,3	82,4	84,3
10. TC7103	78,4	78,4	80,4	80,4	80,4	78,4	80,4	80,4	78,4	78,4	80,4	78,4	80,4	80,4	80,4	84,3	82,4	84,3
11. TA37666	78,4	78,4	80,4	80,4	80,4	78,4	80,4	80,4	78,4	78,4	80,4	78,4	80,4	80,4	80,4	80,4	78,4	80,4

12. TC67603	80.4	80.4	78.4	82.4	82.4	80.4	82.4	82.4	78.4	80.4	82.4	80.4	82.4	82.4	82.4	82.4	80.4	82.4
13. 0s06g08500	70.6	70.6	72.5	72.5	72.5	70.6	72.5	72.5	70.6	70.6	72.5	70.6	72.5	72.5	72.5	72.5	70.6	72.5
14. 0s02g55250	82.4	82.4	80.4	84.3	84.3	82.4	84.3	84.3	80.4	82.4	84.3	82.4	84.3	84.3	84.3	84.3	82.4	84.3
15. 0s09g25040	74.1	74.1	72.2	75.9	75.9	74.1	75.9	75.9	72.2	74.1	75.9	75.9	75.9	75.9	75.9	75.9	74.1	75.9
16. TA2544	80.4	80.4	78.4	82.4	82.4	80.4	82.4	82.4	78.4	80.4	82.4	80.4	82.4	82.4	82.4	82.4	82.4	82.4
17. DV982110	82.4	82.4	80.4	84.3	84.3	82.4	84.3	84.3	80.4	82.4	84.3	82.4	84.3	84.3	84.3	84.3	82.4	84.3
18. WO051050segID516	82.4	82.4	80.4	84.3	84.3	82.4	84.3	84.3	80.4	82.4	84.3	82.4	84.3	84.3	84.3	84.3	82.4	84.3
19. AT4G02590		100.0	98.0	96.1	96.1	92.2	94.1	96.1	92.2	92.2	94.1	94.1	96.1	96.1	96.1	96.1	92.2	94.1
20. TC10015_part	100.0		98.0	96.1	96.1	92.2	94.1	96.1	92.2	92.2	94.1	94.1	96.1	96.1	96.1	96.1	92.2	94.1
21. AT1G03040	100.0	100.0		94.1	94.1	90.2	92.2	94.1	92.2	90.2	92.2	92.2	94.1	96.1	96.1	96.1	92.2	94.1
22. TA13791	100.0	100.0	100.0		100.0	96.1	98.0	100.0	96.1	96.1	98.0	98.0	100.0	100.0	100.0	100.0	90.2	92.2
23. TC60118	100.0	100.0	100.0	100.0		96.1	98.0	100.0	96.1	96.1	98.0	98.0	100.0	100.0	100.0	100.0	96.1	98.0
24. TC60119	96.1	96.1	96.1	96.1	96.1		98.0	96.1	92.2	94.1	96.1	94.1	96.1	96.1	96.1	96.1	96.1	98.0
25. TC61833	98.0	98.0	98.0	98.0	98.0	98.0		98.0	94.1	96.1	98.0	96.1	98.0	98.0	98.0	98.0	98.0	100.0
26. TA3392	100.0	100.0	100.0	100.0	100.0	96.1	98.0		96.1	96.1	98.0	98.0	100.0	100.0	100.0	100.0	96.1	98.0
27. WO051050segID1671	98.0	98.0	98.0	98.0	98.0	94.1	96.1	98.0		92.2	94.1	94.1	96.1	96.1	96.1	96.1	92.2	94.1
28. Pt_TC63334	96.1	96.1	96.1	96.1	96.1	94.1	96.1	96.1	94.1		98.0	94.1	96.1	96.1	96.1	96.1	94.1	96.1
29. TA3263	98.0	98.0	98.0	98.0	98.0	96.1	98.0	98.0	96.1	98.0		96.1	98.0	98.0	98.0	98.0	96.1	98.0
30. Pt_scaff_II.416	100.0	100.0	100.0	100.0	100.0	96.1	98.0	100.0	98.0	96.1	98.0		98.0	98.0	98.0	98.0	94.1	96.1
31. TA2825	100.0	100.0	100.0	100.0	100.0	96.1	98.0	100.0	98.0	96.1	98.0	100.0		100.0	100.0	100.0	96.1	98.0
32. TC207545	100.0	100.0	100.0	100.0	100.0	96.1	98.0	100.0	98.0	96.1	98.0	100.0	100.0		100.0	100.0	96.1	98.0
33. TC229602	100.0	100.0	100.0	100.0	100.0	96.1	98.0	100.0	98.0	96.1	98.0	100.0	100.0	100.0		100.0	96.1	98.0
34. TA46156	100.0	100.0	100.0	100.0	100.0	96.1	98.0	100.0	98.0	96.1	98.0	100.0	100.0	100.0	100.0		96.1	98.0
35. AK247217	98.0	98.0	98.0	98.0	98.0	98.0	100.0	98.0	96.1	96.1	98.0	98.0	98.0	98.0	98.0	98.0	96.1	98.0
36. TA30646	98.0	98.0	98.0	98.0	98.0	98.0	100.0	98.0	96.1	96.1	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0
37. TC8633	96.1	96.1	96.1	98.0	98.0	98.0	100.0	98.0	96.1	96.1	98.0	98.0	98.0	98.0	98.0	98.0	100.0	
38. TC15501	98.0	98.0	98.0	98.0	98.0	94.1	96.1	98.0	96.1	94.1	96.1	98.0	98.0	98.0	98.0	98.0	100.0	100.0
39. TA1140	98.0	98.0	98.0	98.0	98.0	94.1	96.1	98.0	96.1	94.1	96.1	98.0	98.0	98.0	98.0	98.0	96.1	96.1
40. TA3490_part	98.0	98.0	98.0	98.0	98.0	94.1	96.1	98.0	96.1	94.1	96.1	98.0	98.0	98.0	98.0	98.0	96.1	96.1
41. CK293938	98.0	98.0	98.0	98.0	98.0	96.1	96.1	98.0	96.1	94.1	96.1	98.0	98.0	98.0	98.0	98.0	96.1	96.1
42. TC12771	98.0	98.0	98.0	98.0	98.0	96.1	96.1	98.0	96.1	94.1	96.1	98.0	98.0	98.0	98.0	98.0	96.1	96.1
43. TC172581	98.0	98.0	98.0	98.0	98.0	96.1	96.1	98.0	96.1	94.1	96.1	98.0	98.0	98.0	98.0	98.0	96.1	96.1
44. TA28621	98.0	98.0	98.0	98.0	98.0	96.1	96.1	98.0	96.1	94.1	96.1	98.0	98.0	98.0	98.0	98.0	96.1	96.1
45. TA21665	98.0	98.0	98.0	98.0	98.0	96.1	96.1	98.0	96.1	94.1	96.1	98.0	98.9	98.0	98.0	98.0	96.1	96.1
												98.0	98.0	98.0	98.0	98.0	96.1	96.1

46. TA46194	98,0	98,0	98,0	98,0	98,0	94,1	96,1	98,0	96,1	94,1	96,1	98,0	98,0	98,0	98,0	98,0	96,1	96,1
47. DY268946	98,0	98,0	98,0	98,0	98,0	94,1	96,1	98,0	96,1	94,1	96,1	98,0	98,0	98,0	98,0	98,0	96,1	96,1
48. CO123623	98,0	98,0	98,0	98,0	98,0	94,1	96,1	98,0	96,1	94,1	96,1	98,0	98,0	98,0	98,0	98,0	96,1	96,1
49. Pt_scaff_70.65	94,1	94,1	94,1	94,1	94,1	94,1	92,2	94,1	92,2	90,2	92,2	94,1	94,1	94,1	94,1	94,1	92,2	92,2
50. Pt_scaff_XI11.403	94,1	94,1	94,1	94,1	94,1	94,1	92,2	94,1	92,2	90,2	92,2	94,1	94,1	94,1	94,1	94,1	92,2	92,2
51. TA1616	98,0	98,0	98,0	98,0	98,0	94,1	96,1	98,0	96,1	94,1	96,1	98,0	98,0	98,0	98,0	98,0	96,1	96,1
52. TA55042	100,0	100,0	100,0	100,0	100,0	96,1	98,0	100,0	98,0	96,1	98,0	100,0	100,0	100,0	100,0	100,0	98,0	98,0
53. WO051050segID77	96,1	96,1	96,1	96,1	96,1	94,1	94,1	96,1	94,1	92,2	94,1	96,1	96,1	96,1	96,1	96,1	94,1	94,1
54. TC3698	96,1	96,1	96,1	96,1	96,1	96,1	98,0	96,1	94,1	94,1	96,1	96,1	96,1	96,1	96,1	96,1	98,0	98,0
55. 0s03g58330	96,1	96,1	96,1	98,0	98,0	94,1	96,1	98,0	96,1	94,1	96,1	98,0	98,0	98,0	98,0	98,0	96,1	96,1
56. TA139285	96,1	96,1	96,1	98,0	98,0	94,1	96,1	98,0	96,1	94,1	96,1	98,0	98,0	98,0	98,0	98,0	96,1	96,1
57. TabHLH11	94,1	94,1	94,1	96,1	96,1	96,1	98,0	96,1	94,1	94,1	96,1	96,1	96,1	96,1	96,1	96,1	98,0	98,0
58. TC104646	96,1	96,1	96,1	98,0	98,0	94,1	96,1	98,0	96,1	94,1	96,1	98,0	98,0	98,0	98,0	98,0	96,1	96,1
59. TAI26400	96,1	96,1	96,1	98,0	98,0	94,1	96,1	98,0	96,1	94,1	96,1	98,0	98,0	98,0	98,0	98,0	96,1	96,1
60. 0s07g08440	96,1	96,1	96,1	98,0	98,0	94,1	96,1	98,0	96,1	94,1	96,1	98,0	98,0	98,0	98,0	98,0	96,1	96,1
61. 0s02g35660	94,1	94,1	94,1	96,1	96,1	94,1	94,1	96,1	96,1	94,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	94,1
62. DT543504	92,2	92,2	94,1	94,1	94,1	90,2	92,2	94,1	94,1	92,2	94,1	94,1	94,1	94,1	94,1	94,1	92,2	92,2
63. Pt_scaff28.86	92,2	92,2	94,1	94,1	94,1	90,2	92,2	94,1	94,1	92,2	94,1	94,1	94,1	94,1	94,1	94,1	92,2	92,2
64. GSVIVT00016367001	94,1	94,1	96,1	96,1	96,1	92,2	94,1	96,1	96,1	94,1	96,1	96,1	96,1	96,1	96,1	96,1	94,1	94,1
65. TC19278	66,7	66,7	68,6	68,6	68,6	66,7	66,7	68,6	68,6	66,7	68,6	68,6	68,6	68,6	68,6	68,6	66,7	66,7
66. TA14134	92,2	92,2	92,2	94,1	94,1	90,2	92,2	94,1	92,2	92,2	94,1	94,1	94,1	94,1	94,1	94,1	92,2	92,2
67. TC68930	96,1	96,1	96,1	98,0	98,0	94,1	96,1	98,0	96,1	96,1	98,0	98,0	98,0	98,0	98,0	98,0	96,1	96,1
68. TC140470	92,2	92,2	94,1	94,1	94,1	92,2	92,2	94,1	94,1	92,2	94,1	94,1	94,1	94,1	94,1	94,1	92,2	92,2
69. TC253044	92,2	92,2	94,1	94,1	94,1	92,2	92,2	94,1	94,1	92,2	94,1	94,1	94,1	94,1	94,1	94,1	92,2	92,2
70. 0s06g09370	90,2	90,2	92,2	92,2	92,2	90,2	90,2	92,2	92,2	90,2	92,2	92,2	92,2	92,2	92,2	92,2	90,2	90,2

	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
1. AT2G24260	86,3	88,2	88,2	88,2	84,3	84,3	84,3	84,3	84,3	88,2	84,3	88,2	82,4	80,4	86,3	88,2	86,3	82,4
2. AT4G30980	84,3	86,3	86,3	86,3	82,4	82,4	82,4	82,4	82,4	86,3	82,4	86,3	80,4	78,4	84,3	86,3	84,3	80,4
3. AT5G58010	80,4	82,4	82,4	82,4	78,4	78,4	78,4	78,4	78,4	82,4	78,4	82,4	76,5	74,5	80,4	82,4	80,4	82,4
4. TAI2416	86,3	88,2	88,2	88,2	84,3	84,3	84,3	84,3	84,3	88,2	84,3	88,2	82,4	80,4	88,2	88,2	86,3	82,4
5. TA5414	86,3	88,2	88,2	88,2	84,3	84,3	84,3	84,3	84,3	88,2	84,3	88,2	82,4	80,4	88,2	88,2	86,3	82,4
6. TC205173	86,3	88,2	88,2	88,2	84,3	84,3	84,3	84,3	84,3	88,2	84,3	88,2	82,4	80,4	88,2	88,2	86,3	82,4
7. GSVIVT00017237001	86,3	88,2	88,2	88,2	84,3	84,3	84,3	84,3	84,3	88,2	84,3	88,2	82,4	80,4	88,2	88,2	86,3	82,4

8. TC7102	86,3	88,2	88,2	88,2	84,3	84,3	84,3	84,3	84,3	88,2	84,3	88,2	82,4	80,4	88,2	88,2	86,3	82,4
9. TA34455	86,3	88,2	88,2	88,2	84,3	84,3	84,3	84,3	84,3	88,2	84,3	88,2	82,4	80,4	88,2	88,2	86,3	82,4
10. TC7103	82,4	84,3	84,3	84,3	80,4	80,4	80,4	80,4	80,4	84,3	80,4	84,3	78,4	76,5	84,3	84,3	82,4	78,4
11. TA37666	82,4	84,3	84,3	84,3	80,4	80,4	80,4	80,4	80,4	84,3	80,4	84,3	78,4	76,5	84,3	84,3	82,4	78,4
12. TC67603	84,3	86,3	86,3	86,3	82,4	82,4	82,4	82,4	82,4	86,3	82,4	86,3	84,3	82,4	86,3	86,3	84,3	80,4
13. 0s06g08500	74,5	76,5	76,5	76,5	72,5	72,5	72,5	72,5	72,5	76,5	72,5	76,5	70,6	68,6	76,5	76,5	74,5	70,6
14. 0s02g55250	86,3	88,2	88,2	88,2	84,3	84,3	84,3	84,3	84,3	88,2	84,3	88,2	82,4	80,4	88,2	88,2	86,3	82,4
15. 0s09g25040	77,8	79,6	79,6	79,6	75,9	75,9	75,9	75,9	75,9	79,6	75,9	79,6	74,1	72,2	79,6	79,6	77,8	74,1
16. TA2544	84,3	86,3	86,3	86,3	82,4	82,4	82,4	82,4	82,4	86,3	82,4	86,3	80,4	78,4	86,3	86,3	84,3	80,4
17. DV982110	86,3	90,2	88,2	88,2	84,3	84,3	84,3	84,3	84,3	90,2	84,3	88,2	82,4	80,4	86,3	88,2	86,3	82,4
18. WO051050segID516	86,3	90,2	88,2	88,2	84,3	84,3	84,3	84,3	84,3	90,2	84,3	88,2	82,4	80,4	86,3	88,2	86,3	82,4
19. AT4G02590	94,1	90,2	88,2	88,2	88,2	88,2	88,2	88,2	88,2	90,2	90,2	90,2	88,2	86,3	94,1	90,2	88,2	90,2
20. TC10015_part	94,1	90,2	88,2	88,2	88,2	88,2	88,2	88,2	88,2	90,2	90,2	90,2	88,2	86,3	94,1	90,2	88,2	90,2
21. AT1 G03040	92,2	88,2	86,3	86,3	86,3	86,3	86,3	86,3	86,3	88,2	88,2	88,2	86,3	84,3	92,2	88,2	86,3	88,2
22. TA13791	96,1	94,1	92,2	92,2	92,2	92,2	92,2	92,2	92,2	94,1	94,1	94,1	88,2	86,3	90,2	94,1	92,2	94,1
23. TC60118	96,1	94,1	92,2	92,2	92,2	92,2	92,2	92,2	92,2	94,1	94,1	94,1	88,2	86,3	90,2	94,1	92,2	94,1
24. TC60119	94,1	90,2	88,2	88,2	90,2	90,2	90,2	90,2	90,2	90,2	90,2	90,2	88,2	86,3	86,3	90,2	90,2	94,1
25. TC61833	96,1	92,2	90,2	90,2	90,2	90,2	90,2	90,2	90,2	92,2	92,2	92,2	86,3	84,3	88,2	92,2	90,2	96,1
26. TA3392	96,1	94,1	92,2	92,2	92,2	92,2	92,2	92,2	92,2	94,1	94,1	94,1	88,2	86,3	90,2	94,1	92,2	94,1
27. W0051050segI01671	92,2	90,2	88,2	88,2	88,2	88,2	88,2	88,2	88,2	90,2	90,2	90,2	84,3	82,4	86,3	90,2	88,2	90,2
28. Pt_TC63334	94,1	90,2	88,2	88,2	88,2	88,2	88,2	88,2	88,2	90,2	90,2	90,2	84,3	82,4	86,3	90,2	88,2	92,2
29. TA3263	96,1	92,2	90,2	90,2	90,2	90,2	90,2	90,2	90,2	92,2	92,2	92,2	86,3	84,3	88,2	92,2	90,2	94,1
30. Pt_scaff_II.416	94,1	92,2	90,2	90,2	90,2	90,2	90,2	90,2	90,2	92,2	92,2	92,2	86,3	84,3	88,2	92,2	90,2	92,2
31. TA2825	96,1	94,1	92,2	92,2	92,2	92,2	92,2	92,2	92,2	94,1	94,1	94,1	88,2	86,3	90,2	94,1	92,2	94,1
32. TC207545	96,1	94,1	92,2	92,2	92,2	92,2	92,2	92,2	92,2	94,1	94,1	94,1	88,2	86,3	90,2	94,1	92,2	94,1
33. TC229602	96,1	94,1	92,2	92,2	92,2	92,2	92,2	92,2	92,2	94,1	94,1	94,1	88,2	86,3	90,2	94,1	92,2	94,1
34. TA46156	96,1	94,1	92,2	92,2	92,2	92,2	92,2	92,2	92,2	94,1	94,1	94,1	88,2	86,3	90,2	94,1	92,2	94,1
35. AK247217	94,1	90,2	88,2	88,2	88,2	88,2	88,2	88,2	88,2	90,2	90,2	90,2	84,3	82,4	86,3	90,2	88,2	94,1
36. TA30646	96,1	92,2	90,2	90,2	90,2	90,2	90,2	90,2	90,2	92,2	92,2	92,2	86,3	84,3	88,2	92,2	90,2	96,1
37. TC8633		90,2	88,2	88,2	88,2	88,2	88,2	88,2	88,2	90,2	90,2	90,2	86,3	84,3	88,2	90,2	88,2	92,2
38. TC15501	96,1		92,2	92,2	92,2	92,2	92,2	92,2	92,2	100,0	94,1	98,0	90,2	88,2	94,1	98,0	96,1	90,2
39. TA1140	96,1	98,0		100,0	92,2	92,2	92,2	92,2	92,2	92,2	94,1	94,1	86,3	84,3	88,2	92,2	92,2	88,2
40. TA3490_part	96,1	98,0	100,0		92,2	92,2	92,2	92,2	92,2	92,2	94,1	94,1	86,3	84,3	88,2	92,2	92,2	88,2
41. CK293938	96,1	98,0	100,0	100,0		100,0	100,0	100,0	100,0	92,2	90,2	94,1	90,2	88,2	88,2	92,2	94,1	88,2

42. TC12771	96,1	98,0	100,0	100,0	100,0		100,0	100,0	100,0	92,2	90,2	94,1	90,2	88,2	88,2	92,2	94,1	88,2
43. TC172581	96,1	98,0	100,0	100,0	100,0	100,0		100,0	100,0	92,2	90,2	94,1	90,2	88,2	88,2	92,2	94,1	88,2
44. TA28621	96,1	98,0	100,0	100,0	100,0	100,0	100,0		100,0	92,2	90,2	94,1	90,2	88,2	88,2	92,2	94,1	88,2
45. TA21665	96,1	98,0	100,0	100,0	100,0	100,0	100,0	100,0		92,2	90,2	94,1	90,2	88,2	88,2	92,2	94,1	88,2
46. TA46194	96,1	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0		94,1	98,0	90,2	88,2	94,1	98,0	96,1	90,2
47. DY268946	96,1	98,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	98,0		96,1	88,2	86,3	90,2	94,1	94,1	90,2
48. CO123623	96,1	98,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	98,0	100,0		92,2	90,2	94,1	98,0	98,0	90,2
49. Pt_scaff_70.65	90,2	94,1	96,1	96,1	98,0	98,0	98,0	98,0	98,0	94,1	96,1	96,1		98,0	90,2	90,2	92,2	84,3
50. Pt_scaff_XIII.403	90,2	94,1	96,1	96,1	98,0	98,0	98,0	98,0	98,0	94,1	96,1	96,1	100,0		88,2	88,2	90,2	82,4
51. TA1616	94,1	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	100,0	98,0	98,0	94,1	94,1		94,1	92,2	86,3
52. TA55042	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	94,1	94,1	98,0		96,1	90,2
53. WO051050segID77	94,1	96,1	98,0	98,0	100,0	100,0	100,0	100,0	100,0	96,1	98,0	98,0	96,1	96,1	96,1	96,1		88,2
54. TC3698	98,0	98,0	96,1	96,1	96,1	96,1	96,1	96,1	96,1	98,0	96,1	96,1	92,2	92,2	96,1	96,1	94,1	
55. 0s03g58330	96,1	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	100,0	98,0	98,0	92,2	92,2	96,1	98,0	96,1	98,0
56. TA139285	96,1	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	100,0	98,0	98,0	92,2	92,2	96,1	98,0	96,1	98,0
57. TabHLH11	98,0	98,0	96,1	96,1	96,1	96,1	96,1	96,1	96,1	98,0	96,1	96,1	90,2	90,2	94,1	96,1	94,1	100,0
58. TC104646	96,1	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	100,0	98,0	98,0	92,2	92,2	96,1	98,0	96,1	98,0
59. TAI26400	96,1	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	100,0	98,0	98,0	92,2	92,2	96,1	98,0	96,1	98,0
60. 0s07g08440	96,1	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	100,0	98,0	98,0	92,2	92,2	96,1	98,0	96,1	98,0
61. 0s02g35660	94,1	98,0	96,1	96,1	96,1	96,1	96,1	96,1	96,1	98,0	96,1	96,1	92,2	92,2	96,1	96,1	96,1	96,1
62. DT543504	92,2	96,1	94,1	94,1	94,1	94,1	94,1	94,1	94,1	96,1	94,1	94,1	88,2	88,2	94,1	94,1	92,2	94,1
63. Pt_scaff28.86	92,2	96,1	94,1	94,1	94,1	94,1	94,1	94,1	94,1	96,1	94,1	94,1	88,2	88,2	94,1	94,1	92,2	94,1
64. GSVIVT00016367001	94,1	98,0	96,1	96,1	96,1	96,1	96,1	96,1	96,1	98,0	96,1	96,1	90,2	90,2	96,1	96,1	94,1	96,1
65. TC19278	66,7	70,6	68,6	68,6	68,6	68,6	68,6	68,6	68,6	70,6	68,6	68,6	66,7	66,7	68,6	68,6	68,6	68,6
66. TA14134	92,2	96,1	94,1	94,1	94,1	94,1	94,1	94,1	94,1	96,1	92,2	94,1	88,2	88,2	94,1	94,1	92,2	92,2
67. TC68930	96,1	100,0	98,0	98,0	98,0	98,0	98,0	98,0	98,0	100,0	98,0	98,0	92,2	92,2	98,0	98,0	96,1	98,0
68. TC140470	92,2	96,1	94,1	94,1	94,1	94,1	94,1	94,1	94,1	96,1	94,1	94,1	90,2	90,2	94,1	94,1	94,1	94,1
69. TC253044	92,2	96,1	94,1	94,1	94,1	94,1	94,1	94,1	94,1	96,1	94,1	94,1	90,2	90,2	94,1	94,1	94,1	94,1
70. 0s06g09370	90,2	94,1	92,2	92,2	92,2	92,2	92,2	92,2	92,2	94,1	92,2	92,2	88,2	88,2	92,2	92,2	92,2	92,2

	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70
1. AT2G24260	88,2	88,2	88,2	88,2	88,2	86,3	88,2	90,2	90,2	92,2	64,7	88,2	94,1	86,3	86,3	88,2
2. AT4G30980	86,3	86,3	86,3	86,3	86,3	84,3	88,2	90,2	90,2	92,2	64,7	90,2	96,1	86,3	86,3	88,2
3. AT5G58010	84,3	84,3	84,3	84,3	84,3	82,4	84,3	88,2	90,2	92,2	68,6	86,3	96,1	82,4	82,4	84,3

4. TAl2416	88,2	88,2	88,2	88,2	88,2	86,3	88,2	90,2	90,2	92,2	64,7	90,2	94,1	86,3	86,3	88,2
5. TA5414	88,2	88,2	88,2	88,2	88,2	86,3	88,2	90,2	90,2	92,2	64,7	90,2	94,1	86,3	86,3	88,2
6. TC205173	88,2	88,2	88,2	88,2	88,2	86,3	88,2	90,2	90,2	92,2	64,7	90,2	94,1	86,3	86,3	88,2
7. GSVIVT00017237001	88,2	88,2	88,2	88,2	88,2	86,3	88,2	90,2	90,2	92,2	64,7	90,2	94,1	86,3	86,3	88,2
8. TC7102	88,2	88,2	88,2	88,2	88,2	86,3	88,2	90,2	90,2	92,2	64,7	90,2	94,1	86,3	86,3	88,2
9. TA34455	88,2	88,2	88,2	88,2	88,2	86,3	88,2	90,2	90,2	92,2	64,7	90,2	94,1	86,3	86,3	88,2
10. TC7103	84,3	84,3	84,3	84,3	84,3	82,4	86,3	88,2	88,2	90,2	66,7	90,2	94,1	84,3	84,3	86,3
11. TA37666	84,3	84,3	84,3	84,3	84,3	82,4	86,3	88,2	88,2	90,2	66,7	90,2	94,1	84,3	84,3	86,3
12. TC67603	86,3	86,3	86,3	86,3	86,3	84,3	86,3	88,2	88,2	90,2	64,7	88,2	92,2	84,3	84,3	86,3
13. 0s06g08500	76,5	76,5	76,5	76,5	76,5	74,5	78,4	80,4	80,4	82,4	64,7	82,4	86,3	76,5	76,5	78,4
14. 0s02g55250	88,2	88,2	88,2	88,2	88,2	86,3	88,2	90,2	90,2	92,2	64,7	90,2	94,1	86,3	86,3	88,2
15. 0s09g25040	79,6	79,6	79,6	79,6	79,6	77,8	79,6	81,5	81,5	83,3	57,4	81,5	85,2	77,8	77,8	79,6
16. TA2544	86,3	86,3	86,3	86,3	86,3	84,3	86,3	88,2	88,2	90,2	62,7	88,2	92,2	84,3	84,3	86,3
17. DV982110	88,2	88,2	88,2	88,2	88,2	86,3	90,2	92,2	92,2	94,1	66,7	88,2	96,1	88,2	88,2	90,2
18. WO051050segID516	88,2	88,2	88,2	88,2	88,2	86,3	90,2	92,2	92,2	94,1	66,7	88,2	96,1	88,2	88,2	90,2
19. AT4G02590	90,2	90,2	88,2	90,2	90,2	88,2	72,5	74,5	74,5	76,5	49,0	74,5	78,4	70,6	70,6	72,5
20. TC10015_part	90,2	90,2	88,2	90,2	90,2	88,2	72,5	74,5	74,5	76,5	49,0	74,5	78,4	70,6	70,6	72,5
21. AT1 G03040	88,2	88,2	86,3	88,2	88,2	86,3	72,5	74,5	74,5	76,5	49,0	76,5	80,4	70,6	70,6	72,5
22. TA13791	92,2	92,2	90,2	92,2	92,2	90,2	74,5	76,5	76,5	78,4	51,0	76,5	80,4	72,5	72,5	74,5
23. TC60118	92,2	92,2	90,2	92,2	92,2	90,2	74,5	76,5	76,5	78,4	51,0	76,5	80,4	72,5	72,5	74,5
24. TC60119	88,2	88,2	90,2	88,2	88,2	86,3	74,5	74,5	74,5	76,5	51,0	74,5	78,4	72,5	72,5	74,5
25. TC61833	90,2	90,2	92,2	90,2	90,2	88,2	74,5	76,5	76,5	78,4	51,0	76,5	80,4	72,5	72,5	74,5
26. TA3392	92,2	92,2	90,2	92,2	92,2	90,2	74,5	76,5	76,5	78,4	51,0	76,5	80,4	72,5	72,5	74,5
27. WO051050segID1671	88,2	88,2	86,3	88,2	88,2	90,2	72,5	74,5	74,5	76,5	49,0	74,5	78,4	74,5	74,5	74,5
28. Pt_TC63334	88,2	88,2	88,2	88,2	88,2	86,3	72,5	74,5	74,5	76,5	49,0	74,5	78,4	70,6	70,6	72,5
29. TA3263	90,2	90,2	90,2	90,2	90,2	88,2	74,5	76,5	76,5	78,4	51,0	76,5	80,4	72,5	72,5	74,5
30. Pt_scaff_II.416	90,2	90,2	88,2	90,2	90,2	88,2	72,5	74,5	74,5	76,5	49,0	74,5	78,4	70,6	70,6	72,5
31. TA2825	92,2	92,2	90,2	92,2	92,2	90,2	74,5	76,5	76,5	78,4	51,0	76,5	80,4	72,5	72,5	74,5
32. TC207545	92,2	92,2	90,2	92,2	92,2	90,2	74,5	76,5	76,5	78,4	51,0	76,5	80,4	72,5	72,5	74,5
33. TC229602	92,2	92,2	90,2	92,2	92,2	90,2	74,5	76,5	76,5	78,4	51,0	76,5	80,4	72,5	72,5	74,5
34. TA46156	92,2	92,2	90,2	92,2	92,2	90,2	74,5	76,5	76,5	78,4	51,0	76,5	80,4	72,5	72,5	74,5
35. AK247217	88,2	88,2	90,2	88,2	88,2	86,3	72,5	74,5	74,5	76,5	49,0	74,5	78,4	70,6	70,6	72,5
36. TA30646	90,2	90,2	92,2	90,2	90,2	88,2	74,5	76,5	76,5	78,4	51,0	76,5	80,4	72,5	72,5	74,5
37. TC8633	92,2	92,2	92,2	92,2	92,2	90,2	76,5	78,4	78,4	80,4	52,9	78,4	82,4	74,5	74,5	76,5

38. TC15501	92,2	92,2	90,2	92,2	92,2	90,2	80,4	82,4	82,4	84,3	56,9	78,4	86,3	78,4	78,4	80,4
39. TA1140	94,1	94,1	92,2	94,1	94,1	92,2	78,4	80,4	80,4	82,4	54,9	78,4	84,3	76,5	76,5	78,4
40. TA3490_part	94,1	94,1	92,2	94,1	94,1	92,2	78,4	80,4	80,4	82,4	54,9	78,4	84,3	76,5	76,5	78,4
41. CK293938	90,2	90,2	88,2	90,2	90,2	88,2	78,4	76,5	76,5	78,4	52,9	76,5	80,4	76,5	76,5	78,4
42. TC12771	90,2	90,2	88,2	90,2	90,2	88,2	78,4	76,5	76,5	78,4	52,9	76,5	80,4	76,5	76,5	78,4
43. TC172581	90,2	90,2	88,2	90,2	90,2	88,2	78,4	76,5	76,5	78,4	52,9	76,5	80,4	76,5	76,5	78,4
44. TA28621	90,2	90,2	88,2	90,2	90,2	88,2	78,4	76,5	76,5	78,4	52,9	76,5	80,4	76,5	76,5	78,4
45. TA21665	90,2	90,2	88,2	90,2	90,2	88,2	78,4	76,5	76,5	78,4	52,9	76,5	80,4	76,5	76,5	78,4
46. TA46194	92,2	92,2	90,2	92,2	92,2	90,2	80,4	82,4	82,4	84,3	56,9	78,4	86,3	78,4	78,4	80,4
47. DY268946	92,2	92,2	90,2	92,2	92,2	90,2	78,4	76,5	76,5	78,4	52,9	76,5	80,4	72,5	72,5	74,5
48. CO123623	92,2	92,2	90,2	92,2	92,2	90,2	78,4	80,4	80,4	82,4	54,9	78,4	84,3	76,5	76,5	78,4
49. Pt_scaff_70.65	86,3	86,3	84,3	86,3	86,3	84,3	74,5	74,5	74,5	76,5	52,9	74,5	78,4	72,5	72,5	74,5
50. Pt_scaff_XIII.403	84,3	84,3	82,4	84,3	84,3	82,4	72,5	72,5	72,5	74,5	51,0	72,5	76,5	70,6	70,6	72,5
51. TA1616	90,2	90,2	88,2	90,2	90,2	88,2	76,5	78,4	78,4	80,4	52,9	78,4	82,4	74,5	74,5	76,5
52. TA55042	92,2	92,2	90,2	92,2	92,2	90,2	78,4	80,4	80,4	82,4	54,9	78,4	84,3	76,5	76,5	78,4
53. WO051050segID77	90,2	90,2	88,2	90,2	90,2	88,2	78,4	78,4	78,4	80,4	54,9	76,5	82,4	76,5	76,5	78,4
54. TC3698	90,2	90,2	92,2	90,2	90,2	88,2	72,5	76,5	78,4	80,4	52,9	74,5	82,4	70,6	70,6	72,5
55. 0s03g58330		100,0	98,0	100,0	100,0	98,0	78,4	80,4	80,4	82,4	54,9	78,4	84,3	76,5	76,5	78,4
56. TA139285	100,0		98,0	100,0	100,0	98,0	78,4	80,4	80,4	82,4	54,9	78,4	84,3	76,5	76,5	78,4
57. TabHLH11	98,0	98,0		98,0	98,0	96,1	78,4	80,4	80,4	82,4	54,9	78,4	84,3	76,5	76,5	78,4
58. TC104646	100,0	100,0	98,0		100,0	98,0	78,4	80,4	80,4	82,4	54,9	78,4	84,3	76,5	76,5	78,4
59. TA126400	100,0	100,0	98,0	100,0		98,0	78,4	80,4	80,4	82,4	54,9	78,4	84,3	76,5	76,5	78,4
60. 0s07908440	100,0	100,0	98,0	100,0	100,0		76,5	78,4	78,4	80,4	52,9	76,5	82,4	78,4	78,4	78,4
61. 0s02g35660	98,0	98,0	96,1	98,0	98,0	98,0		88,2	88,2	90,2	64,7	82,4	88,2	88,2	88,2	90,2
62. DT543504	96,1	96,1	94,1	96,1	96,1	96,1	98,0		94,1	96,1	66,7	82,4	92,2	88,2	88,2	90,2
63. Pt_scaff28.86	96,1	96,1	94,1	96,1	96,1	96,1	98,0	96,1		98,0	68,6	82,4	94,1	90,2	90,2	92,2
64. GSVIVT00016367001	98,0	98,0	96,1	98,0	98,0	98,0	100,0	98,0	98,0		70,6	84,3	96,1	90,2	90,2	92,2
65. TC19278	70,6	70,6	68,6	70,6	70,6	70,6	72,5	70,6	70,6	72,5		62,7	68,6	62,7	62,7	64,7
66. TA14134	94,1	94,1	92,2	94,1	94,1	94,1	92,2	94,1	94,1	96,1	72,5		88,2	78,4	78,4	80,4
67. TC68930	100,0	100,0	98,0	100,0	100,0	100,0	98,0	98,0	98,0	100,0	72,5	96,1		86,3	86,3	88,2
68. TC140470	96,1	96,1	94,1	96,1	96,1	96,1	98,0	96,1	98,0	98,0	70,6	94,1	98,0		100,0	96,1
69. TC253044	96,1	96,1	94,1	96,1	96,1	96,1	98,0	96,1	98,0	98,0	70,6	94,1	98,0	100,0		96,1
70. 0s06g09370	94,1	94,1	92,2	94,1	94,1	94,1	96,1	94,1	96,1	96,1	68,6	92,2	96,1	98,0	98,0	

Exemplo 46: A identificação dos domínios compreendidos nas sequências de polipeptídeos úteis na realização de métodos da invenção

O recurso integrado das famílias de proteínas, o banco de dados de locais e domínios (InterPro) é uma interface integrada para os bancos de dados assinados comumente usados para as buscas com base no texto e sequência. O banco de dado InterPro combina estes bancos de dados, que usam as metodologias diferentes e variam os graus de informação biológica sobre as proteínas bem caracterizadas para derivar os sinais da proteína. A colaboração dos bancos de dados incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. O Pfam é uma ampla coleção de alinhamentos de sequência múltipla e modelos Markov secretos que cobrem muitos dos domínios de proteína comuns e famílias. O Pfam é hospedado no Sanger Institute server in the United Kingdom. O Interpro é hospedado no European Bioinformatics Institute in the United Kingdom.

Os resultados da avaliação InterPro de uma sequência de polipeptídeo como representado pela SEQ ID N°: 245 são apresentados na Tabela E4.

Tabela E4: Resultados de avaliação InterPro (maiores números de acessão) de uma sequência de polipeptídeo como representado pela SEQ ID N°: 245.

banco de dados	**número de acessão**	**nome de acessão**	**coordenadas de aminoácido na SEQ ID N°: 245; valor E**
InterPro	IPR001092	dimerização da região bHLH de Hélice-Laço-Hélice básico	
HMMPfam	PF00010.14	domínio de ligação de DNA de Hélice-Laço-Hélice	T[127-176] 1.6e-06
HMMSmart	SM00353	nenhuma descrição	T[132-181] 1.5e-09
ProfileScan	PS50888	HLH	T[120-176] 12.451
InterPro	IPR011598	ligação de DNA de Hélice-Laço-Hélice	
superfamília	SSF47459	domínio de ligação de DNA de Hélice-Laço-Hélice	T[122-195] 1.8e-14

Exemplo 47: Predição da topologia de uma sequência de polipeptídeos úteis na realização de métodos da invenção

O TargetP 1.1 prediz a localização subcelular das proteínas

eucarióticas. A indicação da localização é com base na presença predita de qualquer uma das pré-sequências do terminal N: peptídeo de transição de cloroplasto (cTP), peptídeo de alvejamento mitocrondrial (mTP) ou peptídeo de sinal de caminho secretório (SP). As contagens em que a predição final não são baseados nas reais probabilidades, e estas não são necessariamente adicionadas a um. Entretanto, a localização com a contagem mais alta é igualmente de acordo com o TargetP, e a conexão entre as contagens (a classe de confiabilidade) pode ser uma indicação de como certas predições são. A classe de confiabilidade (RC) varia de 1 a 5, onde 1 indica a predição mais forte. O TargetP é mantido no servidor da Technical University of Denmark.

Para as sequências preditas por conter uma pré-sequência de terminal N e um local de clivagem potencial também pode ser predito.

Um número de parâmetros foram selecionados, tal como grupo de organismos (vegetal ou não vegetal), séries de cortes (nenhum, série pré-definida de cortes, ou série especificadas pelo uso de cortes), e o cálculo da predição dos locais de clivagem (sim ou não).

Os resultados da análise TargetP 1.1 de uma sequência de polipeptídeo como representado pela SEQ ID N°: 245 são apresentados na Tabela E5. O grupo de organismos "vegetal" foi selecionado, nos cortes definidos, e o comprimento predito do peptídeo de transição requerido. A localização subcelular de uma sequência de polipeptídeo como representado pela SEQ ID N°: 245 pode ser o citoplasma ou núcleo, nenhum peptídeo de transição predito.

Tabela E5: Análise TargetP 1.1 de uma sequência de polipeptídeo como representado pela SEQ ID N°: 245

Comprimento (AA)	281
Peptídeo de transição cloroplástica	0,094
Peptídeo de transição mitocondrial	0,166
Peptídeo de sinal do caminho secretório	0,034
Outro alvejamento subcelular	0,855
Localização predita	/
Classe de confiabilidade	2
Comprimento de peptídeo de transição predita	/

Quando analisado com PLOC (Park and Kanehisa, Bioinformatics, 19, 1656-1663, 2003) ou com PSORT (URL: psort.org), a proteína é predita ter uma localização nuclear.

Muitos outros algorítmos podem ser usados para realizar tais análises, incluindo:

- ChloroP 1.1 hospedado no servidor da Technical University of Denmark;
- Protein Prowler Subcellular Localisation Predictor versão 1.2 hospedado no servidor do Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia;
- PENCE Proteome Analyst PA-GOSUB 2.5 hospedado no servidor da University of Alberta edmonton, Alberta, Canada;
- TMHMM, hospedado no servidor da Technical University of Denmark;

Exemplo 48: Ensaio funcional para o polipeptídeo semelhante ao bHLH11

Um ensaio de ligação de DNA para AtbHLH6 é fornecido em Dombrecht et al. (2007), este método pode ser usado por qualquer proteína semelhante à bHLH11. Brevemente, um fragmento 1000-bp da sequência codificadora AtbHLH6 abrange os códons 285 a 623 foi amplificada a partir do DNA genômico e clonado no vetor pTacLCELD6.His digerido por NheI-BamHI (Xue, Plant J. 41, 638-649, 2005). A construção resultante codificado pelos últimos 338 aminoácidos de AtbHLH6 (incluindo a região bHLH) na estrutura com uma proteína repórter CeID e um rótulo 6xHis. A amplificação e clonagem correta foram verificadas pelo sequenciamento de DNA.

A determinação da sequência de consenso do motivo de ligação de DNA MYC2 e a afinidade de ligação relativa destes locais foi feito de acordo com o Xue (2005), em que uma fusão de uma proteína de ligação de DNA (DBP) à celulase 6-His-tagged D (CELD) serve ambos para os meios pela purificação por afinidade do complexo DBP-DNA na seleção dos locais

de ligação a partir do grupo de oligonucleotídeos da sequência aleatória biontinilada e como um repórter para a medição da atividade de ligação de DNA.

Para a seleção de locais de ligação usando o papel celulose como uma matriz de afinidade, a DBP-CELD foi incubada em temperatura ambiente por 1 hora com 20 ng de um Bio-RS-Oligo rotulado por biotina em 40 µl do tampão de ligação/lavagem (descrito cima) contendo 1 mM de EDTA, 0,25 µg μl^{-1} poli d(AC-TG), 1 mg de ml^{-1} de BSA e 10 % de glicerol. Uma quantidade apropriada de DBP-CELD bruto usado para a seleção do local de ligação foi o que atingiu 20 a 30 % da eficiência da ligação de celulose relativa (porcentagem de ligação da atividade celulose à lavagem após a celulose). As misturas DBP-CELD/Bio-RS-Oligo foram transferidss aos reservatórios de microplacas de 96 reservatórios contendo papel de filtro Whatman 1 (4 mm^2) que foi pré-embebido com 10 µl da solução de bloqueio [tampão ligação/lavagem contendo 0,5 µg μl^{-1} poli d(AC-TG) e 10 % de glicerol]. Após a incubação a 0° C por 1 hora com agitação moderada, o papel celulose foi lavado seis vezes com o tampão ligação/lavagem contendo 1 mM de EDTA e 0,1 mg ml^{-1} de BSA. O uso da lavagem extensiva na seleção do alvo local foi baseado nas observações da estabilidade relativamente alta do complexo de oligonucleotídeo DBP imobilizado me matriz sólida. Os locais de ligação alvo realizaram DBP-CELD foram eluídos a 40° C por 15 minutos com 40 µl de celulase eluindo-se o tampão [10 mM de HEPES, pH7, 50 mM de KCl, 0,2 mM de EDTA, 4 mM de celobiose, 0,05 mg ml^{-1} de BSA e 10 µg ml^{-1} de um iniciador específico da sequência sentido SP-S]. O eluído foi usado pela amplificação PCR dos oligonucleotídeos selecionados. O produto PCR (0,1 de µl) sem purificação foi usado para o próximo ciclo de seleção do local.

Alternativamente, 6-His tagged DBP-CELD (10 a 25 µg de proteína bruta) foi incubada em 60 µl de tampão PNT (50 mM de fosfato de

sódio, pH 8.0, 300 mM de NaCl e 0,05 % de Tween 20) contendo 10 mM de imidazol e 350 μg de pérolas de agarose magnéticas Ni-NTA (Qiagen) em temperatura ambiente por 45 a 60 minutos com a agitação moderada em um misturados de microtubo (Tomy Seiko Co., Tokyo, Japan). As perólas magnéticas Ni-NTA foram coletadas no local do tubo pela posição dos tubos em um magneto de 12 tubos (Qiagen). Proteínas de não ligação foram removidas pela lavagem duas vezes com 150 a 200 μl de PNT contendo 20 mM de imidazol e uma vez com 60 μl do tampão ligação/lavagem contendo 2 mM de $MgCl_2$, 1 mg ml^{-1} de BSA e 10 % de glicerol. As perólas lavadas foram recolocadas em suspensão em 40 μl do tampão ligação/lavagem contendo 2 mM de $MgCl_2$, 0,25 μg μl^{-1} poli d(AC-TG), 1 mg ml^{-1} de BSA, 0,0025 % de Nonidet P-40, 10 % de glicerol e oligonucleotídeos rotulados por biotina (50 ng de Bio-RS-Oligo ou 1 μl do produto PCR amplificado por oligonucleotídeos previamente selecionados). A suspensão foi incubada em temperatura ambiente por 1,5 horas com a agitação moderada no misturado de microtubo. Após a lavagem, quatro vezes (3 a 4 minutos cada lavagem) com 150 a 200 μl de tampão ligação/lavagem contendo 2 mM de $MgCl_2$, 0,1 mg ml^{-1} de BSA e 0,0025 % de Nonidet P-40, as pérolas realizaram a ligação de local alvo BDP-CELD r foram recolocadas em suspensão em 8 μl de 5mM de Tris-Cl (pH 8.0)/0,5 mM de EDTA contendo 5 μg ml^{-1} de um iniciador específico da sequência sentido (SP-S) e a suspensão foi transferida ao tubo limpo e usada para a amplificação de PCR dos oligonucleotídeos selecionados. O produto PCR (1 μl) sem purificação foi usado para o próximo ciclo de seleção do local.

Para os ensaios EMSA, as proteínas 6-His-tagged DBP-CELD foram purificadas usando as pérolas de agarose magnéticas Ni-NTA usando um alto tampão de ligação estringente (50 mM de fosfato de sódio, pH8.0, 1M de NaCl, 10 % de glicerol, 1 % de Tween 20 e 10 mM de imidazol) e o restante do procedimento seguindo as instruções dos fabricantes.

Oligonucleotídeos sintéticos filamentados duplos (30 a 45 fmol) rotulados com digoxigenina na extremida final 3' foram foram incubados com um DBP purificado (30 a 75 ng) em 15 µl do tampão de ligação [25 mM de HEPES/KOH, pH 7.0, 50 mM de KCl, 0,5 mM de DTT, 2 mM de $MgCl_2$, 0,2 µg $µl^{-1}$ poli d(AC-TG), 0,3 mg m^{-1} de BSA e 10 % de glicerol]. Após a incubação em temperatura ambiente por 30 minutos, os complexos DBP/DNA foram separados a partir de sondas livres em um gel de poliacrilamida a 6 % em um tampão de Tris-acetato a 40-mM (pH 7.5) contendo 5 mM de acetato de Na, 0,5 mM de EDTA e 5 % de glicerol. Os complexos DBP/DNA e as sondas livres nos géis após a eletroforese foram transferidos a uma membrana Hybond Np. O anticorpo anti-dioxigenina conjugado por fosfatase alcalina e um substrato quimioluminescente, CDP-Star (Roche Diagnostics) foi usado para a detecção de digoxigenina de acordo com as instruções do fabricante.

Detalhes adicionais são fornecidos em Xue (2005).

Exemplo 49: Clonagem da sequência do ácido nucleico usado nos métodos da invenção

A sequência do ácido nucleico usado nos métodos da invenção e que compreendem a SEQ ID N°: 244 foi amplificado por PCR usando como modelo uma feitas comumente da biblioteca de cDNA de mudas *Triticum aestivum* (em pCMV Sport 6.0; Invitrogen, Paisley, UK). O PCR foi realizado usando polimerase de DNA Hifi Taq em condições padrões, usando 200 ng do modelo em uma mistura de PCR a 50 µl. Os iniciadores usados foram prm009718 (SEQ ID N°: 254; sentido, códon de partida em negrito):

5'-ggggacaagtttgtacaaaaaagcaggcttaaacaatggcggggcanccg-3'

e prm009719 (SEQ ID N°: 255; reverso, complementar):

5'-ggggaccactttgtacaagaaagctgggtatgggtgttgcagctgctgtt-3',

que incluem os locais AttB para a recombinação Gateway. O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então

realizada, durante o qual fragmento PCR recombina in vivo com o plasmídio pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada", semelhante ao pbHLH11. O plasmídio pDONR201 foi adquirido de Invitrogen, como partes da tecnologia Gateway®.

O clone de entrada que compreende a SEQ ID N°: 244 foi então usado em uma reação LR com um vetor de destino usado para a transformação de *Oryza sativa*. Este vetor contido como os elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; um cassete de expressão marcador avaliável; e um cassete Gateway pretendido pela recombinação LR *in vivo* com uma sequência do ácido nucleico de interesse já clonado no clone de entrada. Um promotor GOS2 de arroz (SEQ ID N°: 256) para a expressão constitutiva foi localizada a montante deste cassete Gateway.

Após a etapa de recombinação LR, o vetor de expressão resultante pGOS2::semelhante à bHLH11 (Figura 23) foi transformado na cepa da *Agrobacterium* LBA4044 de acordo com os métodos bem conhecidos na técnica.

Exemplo 50: Transformação da planta

A transformação das plantas foi realizado de acordo com os procedimentos resumidos no Exemplo 7.

Exemplo 51: Procedimento de avaliação fenotípica

51.1 Apresentação da avaliação

Aproximadamente 35 transformantes de arroz TO independentes foram gerados. Os transformantes primários foram transferidos a partir de uma câmera de cultura de tecido de uma estufa para o desenvolvimento e coletado da semente T1. Seis eventos, de que uma progênie T1 segregada 3:1 para a presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 mudas T1 contendo o transgene (hetero e homozigotos) e aproximadamente 10 mudas

T1 perderam o transgene (nulizigotos) foram selecionados pela expressão marcadora visual do monitoramento.As plantas transgênicas e os nulizigotos correspondentes foram desenvolvidos em posições aleatórias lado a lado. As condições da estufa foram de dias curtos (12 horas de luz), 28° C na luz e22° C no escuro, e uma umidade relativa de 70 %. O desenvolvimento das plantas sob as condições de não tensão são submetidas à água em intervalos regulares para garantir que a água e nutrientes não são limitantes para satisfazer a necessidade das plantas para completar o crescimento e desenvolvimento.

A partir do estágio de semeadura até o estágio de maturidade das plantas foram passados diversas vezes através de um gabinete de imagem digital. Em cada período de tempo imagens digitais (2048x1536 píxeis, 16 milhões de cores) foram tiradas de cada planta em pelo menos 6 ângulos diferentes.

Avaliação da aridez

As plantas das sementes T2 são desenvolvidas no solo em potes sob condições normais até o método do estágio da parte superior. Estes são então transferidos à seção "seca" onde a irrigação é retida. As sondas de umidade são inseridas em potes escolhidos aleatoriamente para monitorar o teor de água no solo (SWC). Quando SWC estão abaixo de certos limiares, as plantas são automaticamente novamente submetidas à água continuamente até um nível normal ser atingido novamente. As plantas são então re-transferidas novamente a condições normais. O resto do cultivo (amadurecimento da planta, coleta da semente) é o mesmo como para as plantas não desenvolvidas sob as condições de tensão abiótica. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais.

Avaliação da eficiência do uso de nitrogênio

As plantas de arroz das sementes T2 são desenvolvidas no solo em potes sob condições normais exceto para a solução do nutriente. Os potes

são submetidos à água a partir da transplantação do amadurecimento com uma solução de nutriente específico contendo teor de nitrogênio N reduzido (N), usualmente entre 7 a 8 vezes menor. O resto do cultivo (amadurecimento da planta, coleta da semente) é o mesmo como para as plantas não desenvolvidas sob a tensão abiótica. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais.

Avaliação da tensão do sal

As plantas são desenvolvidas em um substrato feito de fibras de coco e argex (razão de 3 para 1). Uma solução de nutriente normal são usados durante as primeiras duas semanas após a transplantação das plantinhas nas estufas. Após as primeiras duas semanas, 25 mM de sal (NaCl) é adicionado à solução de nutriente, até as plantas serem coletadas. Os parâmetros relacionados às sementes são então medidos.

51.2 Análise estatística: Teste F

Dois fatores ANOVA (análise de variantes) foram usados como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformados com o gene da presente invenção. O teste F foi realizado para checar por um efeito do gene em todos os eventos de transformação e para verificar por um total do efeito do gene, também conhecido como um efeito do gene global. O limiar para a significância por um efeito verdadeiro do gene global foi apresentado em uma 5 % do nível de probabilidade para o teste F. Um ponto de valor do teste F significante por um efeito do gene, significando que este não é apenas uma mera presença ou posição do gene que está causando as diferenças no fenotipo.

51.3 Parâmetros medidos

Medição do parâmetro relacionado à biomassa

A partir do estágio de semeadura até o estágio de maturidade das plantas foram passados diversas vezes através de um gabinete de imagem digital. Em cada período de tempo imagens digitais (2048x1536 píxeis, 16 milhões de cores) foram tiradas de cada planta em pelo menos 6 ângulos diferentes.

Uma área da planta acima do solo (ou biomassa da folha) foi determinado pela contagem do número total de píxeis nas imagens digitais a partir das partes de plantas acima do solo discriminadas a partir do fundamento. Este valor foi medido pelas figuras retiradas do mesmo ponto de tempo a partir de ângulos diferentes e foi convertido a um valor de superfície física expressada em quadrado mm pela calibração. Os experimentos mostram que à área da planta acima do solo medida desta maneira correlaciona com a biomassa das partes das plantas trituradas acima. À área triturada acima é a área medida no ponto de tempo em que a planta tem atingido esta biomassa máxima da folha. O vigor precoce é à área das planta (mudas) acima do solo três semanas pós-germinação. O aumento na biomassa da raiz é expressado como um aumento no total da biomassa da raiz (medido como biomassa máxima de raízes observadas durante a vida toda de uma planta); ou como um aumento na raiz/índice de broto (medido como a razão entre massa de raiz e massa de broto no período de desenvolvimento ativo da raiz e broto).

O vigor precoce foi determinado pela contagem do número total de píxeis a partir das partes de plantas acima do solo discriminadas a partir do fundamento. Este valor foi medido pelas figuras retiradas do mesmo ponto de tempo de ângulos diferentes e foi convertido a um valor de superfície física expressada em quadrado mm pela calibração: Os resultados descritos abaixo são para as plantas com três semanas pós-germinação.

Medições dos parâmetros relacionados à semente

As panículas primárias maduras foram coletadas, contadas,

ensacadas, rotuladas com código de barra e então secadas por três dias em um forno a 37° C. As panículas foram então limiadas e todas as sementes foram coletadas e contadas. As cascas enchidas foram separadas a partir de um vazio usando um dispositivo de sopro de ar. As cascas vazias foram descartadas e a fração remanescente foi contada novamente. As cascas enchidas foram pesadas em uma balança analítica. O número de sementes enchidas foi determinado pela contagem do número de cascas enchidas que permanece depois da etapa de separação. A produção total de semente foi medida pela pesagem total das cascas enchidas coletadas a partir da planta. O número total de semente por planta foi medido pela contagem do número de cascas coletadas a partir da planta. Peso de núcleo milhar (TKW) é extrapolado do número de sementes enchidas contadas e seu peso total. O índice de colheita (HI) na presente invenção é definida como a razão entre a produção total de semente e à área triturada acima (mm^2), multiplicada por um fator 10^6. O número total de flores por panícula como definido na presente invenção é a razão entre o número total de sementes e o número de panículas primárias maduras. A taxa de abastecimento da semente como definido na presente invenção é a proporção (expressado como uma %) do número de sementes enchidas no número total de sementes (ou florzinhas).

Exemplo 52: Resultados da avaliação fenotípica das plantas transgênicas

Os resultados da avaliação das plantas transgênicas de arroz que expressam um ácido nucleico semelhante ao bHLH11 sob as condições de não tensão são apresentadas abaixo. Um aumento de mais do que 5 % em pelo menos 2 linhas foi observado pela produção total de semente, número de sementes enchidas, taxa de abastecimento, índice de colheita, número das primeiras panículas e em pelo menos 3 % por peso de núcleo milhar. Os dados no aumento total (medido em todas as linhas do teste) para cada parâmetro são apresentados na Tabela E6:

Tabela E6:

Parâmetro	Aumento total (%)
Peso total das sementes (produção total da semente	20
Número de sementes enchidas	20
Taxa de enchimento	18
Índice de colheita	21

VI. Proteína ASR (ácido abscísico-, tensão-, e maturativo induzido)

Exemplo 53: A identificação das sequências relacionadas à sequência de ácido nucleico ASR usado nos métodos da invenção

As sequências (cDNA eSTs de comprimento total ou genômico) relacionadas a uma sequência do ácido nucleico usadas nosos métodos da presente invenção onde identificados entre aqueles mantidos no banco de dados Entrez Nucleotides at the National Center for Biotechnology Information (NCBI) usando as ferramentas de sequência de banco de dados, tal como a ferramenta de alinhamento local básica (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa foi usado para observar as regiões da similaridade local entre as sequências em comparação com o ácido nucleico ou sequências de polipeptídeos de um banco de dados da sequência e pleo cálculo da significância estatística de pontos. Por exemplo, o polipeptídeo codificado pelo ácido nucleico usado na presente invenção foram usados pelo algorítmo TBLASTN, com ajustes default e o filtro para ignorar baixa complexidade sequência apresentada. A produção da análise foi resumida em comparação aos pares, e classificada de acordo com as contagens de probabilidade (valor E), onde a contagem reflete a probabilidade que um alinhamento particular ocorra pela possibilidade (a parte inferior do valor E, quanto mais significante o acerto). Além disso aos valores E, as comparações também foram contadas pela identidade de porcentagem. A identidade da porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácido nucleico comparado (ou polipeptídeo) em um comprimento particular.

Tabela F1 fornece uma lista de ácido nucleico e sequências de polipeptídeos relacionadas a uma sequência do ácido nucleico usado nos métodos da presente invenção.

Tabela F1: Exemplos de ácidos nucleicos ASR e polipeptídeos.

nome	origem da planta	ácido nucleico da SEQ ID N°:	proteína da SEQ ID N°:
Orysa_ASR_1	*Oryza sativa*	396	397
Zeama_ASR_1	*Zea mays*	401	402
Zeama_ASR_2	*Zea mays*	403	404
Zeama_ASR_3	*Zea mays*	405	406
Zeama_ASR_4	*Zea mays*	407	408
Zeama_ASR_5	*Zea mays*	409	410
Zeama_ASR_6	*Zea mays*	411	412
Zeama_ASR_7	*Zea mays*	413	414
Zeama_ASR_8	*Zea mays*	415	416
Zeama_ASR_9	*Zea mays*	417	418
Lyces_ASR_1	*Lycopersicon esculentum*	419	420
Lyces_ASR_2	*Lilium longiflorum*	421	422
Sacof ASR 1	*Saccharum officinarum*	423	424
Zeama_ASR_10	*Zea mays*	425	426

Exemplo 54: Alinhamento das sequências de polipeptídeos ASR

O alinhamento das sequências de polipeptídeos é realizado usando o programa AlignX a partir da embalagem do vetor NTI (Invitrogen) que é baseado no algorítmo Clustal W popular do alinhamento progressivo (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31:3497-3500). Os valores default são pela penalidade aberta de fenda de 10, pela penalidade de extensão de fenda de 0,1 e a matriz do peso selecionado é Blosum 62 (se os polipeptídeos são alinhados). Uma edição menor do manual pode ser feita ainda para otimizar o alinhamento.

Uma árvore filogenética de polipeptídeos GRP é construído usando um algorítmo de agrupamento próximo como fornecido no programa

AlignX do vetor NTI (Invitrogen).

Exemplo 55: Cálculo da identidade global da porcentagem entre as sequências de polipeptídeos ASR úteis na realização de métodos da invenção

As porcentagens globais de similaridade e identidade entre as sequências de polipeptídeos de comprimento total útil na realização de métodos da invenção são determinadas usando um dos métodos disponíveis na técnica, o software MatGAT (Ferramenta de alinhamento global de matriz) (BMC Bioinformatics. 2003 4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando sequências de DNA ou proteínas. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). O software MatGAT gera as matrizes de similaridade/identidade para sequência de proteínas ou DNA sem necessariamento o pré-alinhamento dos dados. O programa realiza uma série de alinhamento em pares usando o algorítmo de alinhamento global Myers and Miller (com uma penalidade de abertura de fenda de 12, e a penalidade de extensão de fenda de 2), calcula a similaridade e identidade usando por exemplo Blosum 62 (para os polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade da sequência é mostrada na metada da parte inferior da linha de divisão e a identidade de sequência é mostrada na metada da parte superior da linha de divisão diagonal.

Parâmetros usados na comparação foram:

Matriz de contagem:	Blosum62
Primeira fenda:	12
Extensão da fenda:	2

Tabela MATGAT para o alinhamento local de um domínio específico, ou dados em % da identidade/similaridade entre os domínios específicos também podem ser gerados.

Exemplo 56: A identificação dos domínios compreendidos nas sequências de polipeptídeos ASR úteis na realização de métodos da invenção

O recurso integrado das famílias de proteínas, os bancos de dados de locais e domínios (InterPro) é uma interface integrada para os bancos de dados assinados comumente usados para as buscas com base no texto e sequência. O banco de dado InterPro combina estes bancos de dados, que usam as metodologias diferentes e variam os graus de informação biológica sobre as proteínas bem caracterizadas para derivar os sinais da proteína. A colaboração dos bancos de dados incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. O Pfam é uma ampla coleção de alinhamentos de sequência múltipla e modelos Markov secretos que cobrem muitos dos domínios de proteína comuns e famílias. O Pfam é hospedado no servidor Sanger Institute in the United Kingdom. O Interpro é hospedado no European Bioinformatics Institute in the United Kingdom.

As sequências de proteínas que representam o GRP são usadas como questão para buscar o banco de dado InterPro.

Exemplo 57: Predição da topologia das sequências de polipeptídeos ASR úteis na realização de métodos da invenção

O TargetP 1.1 prediz a localização subcelular das proteínas eucarióticas. A indicação da localização é com base na presença predita de qualquer uma das pré-sequências do terminal N: peptídeo de transição de cloroplasto (cTP), peptídeo de alvejamento mitocrondrial (mTP) ou peptídeo de sinal de caminho secretório (SP). As contagens em que a predição final não são baseados nas reais probabilidades, e estes não são necessariamente adicionado a um. Entretanto, a localização com a contagem mais alta é igualmente de acordo com o TargetP, e a conexão entre as contagens (a classe de confiabilidade) pode ser uma indicação de que certas predições são. A classe de confiabilidade (RC) varia de 1 a 5, onde 1 indica a predição mais forte. O TargetP é mantido pela Technical University of Denmark.

Para as sequências preditas por conter uma pré-sequência de

terminal N e um local de clivagem potencial também pode ser predito.

Um número de parâmetros foram selecionados, tal como grupo de organismos (vegetal ou não vegetal), séries de cortes (nenhum, série pré-definida de cortes, ou série especificadas pelo uso de cortes), e o cálculo da predição dos locais de clivagem (sim ou não).

As sequências de proteínas que representam o GRP são usados para questionar o TargetP 1.1. O grupo de organismos "vegetal" é selecionado, nos cortes definidos, e o comprimento predito do peptídeo de transição requerido.

Muitos outros algorítmos podem ser usados para realizar tais análises, incluindo:

- ChloroP 1.1 hospedado no servidor da Technical University of Denmark;
- Protein Prowler Subcellular Localisation Predictor versão 1.2 hospedado no servidor do Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia;
- PENCE Proteome Analyst PA-GOSUB 2.5 hospedado no servidor da University of Alberta edmonton, Alberta, Canada;
- TMHMM, hospedado no servidor da Technical University of Denmark

Exemplo 58: Clonagem de uma sequência de ácido nucleico ASR usado nos métodos da invenção

Clonagem da SEQ ID N°: 396:

A sequência do ácido nucleico SEQ ID N°: 396 usada nos métodos da invenção foi amplificado por PCR usando como modelo uma feitas comumente na biblioteca de cDNA de mudas de *Oryza sativa* (em pCMV Sport 6.0; Invitrogen, Paisley, UK). PCR foi realizado usando a polimerase de DNA Hifi Taq em condições padrões, usando 200 ng do modelo em uma mistura de PCR de 50 µl. Os iniciadores usados foram

prm06442 (SEQ ID N°: 399; sentido, códon de partida em negrito):

5'-ggggacaagtttgtacaaaaaagcaggcttaaacaatggcggaggagaagca-3'

e prm06443 (SEQ ID N°: 400; reverso, complementar):

5'-ggggaccactttgtacaagaaagctgggtcggcgacgttgtgatga-3',

que incluem os locais AttB para a recombinação Gateway. O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante o qual fragmento PCR recombina *in vivo* com o plasmídio pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada". O plasmídio pDONR201 foi adquirido de Invitrogen, como partes da tecnologia Gateway®.

O clone de entrada que compreende SEQ ID N°: 396 foi então usado em uma reação LR com um vetor de destino usado para a transformação de *Oryza sativa*. Este vetor contido como os elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; um cassete de expressão marcador avaliável; e um cassete Gateway pretendido pela recombinação LR *in vivo* com uma sequência do ácido nucleico de interesse já clonado no clone de entrada. Um promotor GOS2 de arroz (SEQ ID N°: 398) para a expessão específica da semente foi localizada a montante deste cassete Gateway.

Após a etapa de recombinação LR, o vetor de expressão resultante pGOS2::GRP (Figura 25) foi transformado na cepa da *Agrobacterium* LBA4044 de acordo com os métodos bem conhecidos na técnica.

Exemplo 59: Transformação da planta

A transformação das plantas foi realizado de acordo com os procedimentos resumidos no Exemplo 7

Exemplo 60: Procedimento de avaliação fenotípica

60.1 Apresentação da avaliação

Aproximadamente 35 transformantes de arroz TO independentes foram gerados. Os primeiros transformantes foram transferidos a partir de uma câmera de cultura de tecido de uma estufa para o desenvolvimento e coletado da semente T1. Seis eventos, de que uma progênie T1 segregada 3:1 para a presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 mudas T1 contendo o transgene (hetero e homozigotos) e aproximadamente 10 mudas T1 perderam o transgene (nulizigotos) foram selecionados pela expressão marcadora visual do monitoramento. As plantas transgênicas e os nulizigotos correspondentes foram desenvolvidos lado a lado em posições aleatórias. As condições da estufa foram de dias curtos (12 horas de luz), 28° C na luz e 22° C no escuro, e uma umidade relativa de 70 %.

Quatro eventos T1 ainda foram avaliados na geração T2 seguindo o mesmo procedimento de avaliação como para a geração T1 mas com mais indivíduos por evento.

A partir do estágio de semeadura até o estágio de maturidade das plantas foram passados diversas vezes através de um gabinete de imagem digital. Em cada período de tempo imagens digitais (2048x1536 píxeis, 16 milhões de cores) foram tiradas de cada planta em pelo menos 6 ângulos diferentes.

Avaliação da aridez

As plantas das sementes T2 são desenvolvidas no solo em potes sob condições normais até estes aproximarem do estágio superior. Estes são então transferidos à seção "seca" onde a irrigação é retida. As sondas de umidade são inseridas em potes escolhidos aleatoriamente para monitorar o teor de água no solo (SWC). Quando SWC estão abaixo de certos limiares, as plantas são automaticamente novamente submetidas à água continuamente até um nível normal ser atingido novamente. As plantas são então re-transferidas novamente a condições normais. O resto do cultivo (amadurecimento da

planta, coleta da semente) é o mesmo como para as plantas não desenvolvidas sob as condições de tensão abiótica. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais. O desenvolvimento das plantas sob as condições de não tensão são fornecidos com água em intervalos regulares para garantir que a água e nutrientes não são limitantes para satisfazer a necessidade das plantas para completar o crescimento e desenvolvimento.

Avaliação da eficiência do uso de nitrogênio

As plantas de arroz das sementes T2 são desenvolvidas no solo em potes sob condições normais exceto para a solução do nutriente. Os potes são submetidos à água a partir da transplantação do amadurecimento com uma solução de nutriente específico contendo teor de nitrogênio N reduzido (N), usualmente entre 7 a 8 vezes menor. O resto do cultivo (amadurecimento da planta, coleta da semente) é o mesmo como para as plantas não desenvolvidas sob a tensão abiótica. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais.

60.2 Análise estatística: Teste F

Dois fatores ANOVA (análise de variantes) foram usados como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformados com o gene da presente invenção. O teste F foi realizado para checar por um efeito do gene em todos os eventos de transformação e para verificar por um total do efeito do gene, também conhecido como um efeito do gene global. O limiar para a significância por um efeito verdadeiro do gene global foi apresentado em uma 5 % do nível de probabilidade para o teste F. Um ponto de valor do teste F significante por um efeito do gene, significando que este não é apenas uma mera presença ou posição do gene que está causando as diferenças no

fenotipo.

Por causa de dois experimentos com os eventos de sobreposição foram realizada, uma análise combinada é realizada. Este é útil para checar a consistência dos efeitos nos dois experimentos, e se este é o caso, para acumular a evidência de ambos experimentos a fim de aumentar a confidência da conclusão. O método usado é um método modelo misturado que leva em conta a estrutura multinível dos dados (isto é experimento - evento - segregantes). Os valores P obtidos em comparação ao teste de razão de probabilidade para distribuições quadradas chi.

60.3 Parâmetros medidos

Medição do parâmetro relacionado à biomassa

Uma área da planta acima do solo (ou biomassa da folha) foi determinado pela contagem do número total de píxeis nas imagens digitais a partir das partes de plantas acima do solo discriminadas a partir do fundamento. Este valor foi medido pelas figuras retiradas do mesmo ponto de tempo a partir de ângulos diferentes e foi convertido a um valor de superfície física expressada em quadrado mm pela calibração. Os experimentos mostram que à área da planta acima do solo medida desta maneira correlaciona com a biomassa das partes das plantas trituradas acima. À área triturada acima é a área medida no ponto de tempo em que a planta tem atingido esta biomassa máxima da folha.

Medições dos parâmetros relacionados à semente

As panículas primárias maduras foram coletadas, contadas, ensacadas, rotuladas com código de barra e então secadas por três dias em um forno a 37° C. As panículas foram então limiadas e todas as sementes foram coletadas e contadas. As cascas enchidas foram separadas a partir de um vazio usando um dispositivo de sopro de ar. As cascas vazias foram descartadas e a fração remanescente foi contada novamente. As cascas enchidas foram pesadas em uma balança analítica. O número de sementes enchidas foi

determinado pela contagem do número de cascas enchidas que permanece depois da etapa de separação. A produção total de semente foi medida pela pesagem total das cascas enchidas coletadas a partir da planta. O número total de semente por planta foi medido pela contagem do número de cascas coletadas a partir da planta. Peso de núcleo milhar (TKW) é extrapolado do número de sementes enchidas contadas e seu peso total. O índice de colheita (HI) na presente invenção é definida como a razão entre a produção total de semente e à área triturada acima (mm^2), multiplicada por um fator 10^6. O número total de flores por panícula como definido na presente invenção é a razão entre o número total de sementes e o número de panículas primárias maduras. A taxa de abastecimento da semente como definido na presente invenção é a proporção (expressado como uma %) do número de sementes enchidas no número total de sementes (ou florzinhas).

Exemplo 61: Resultados da avaliação fenotípica das plantas transgênicas

As plantas transgênicas de arroz que expressam o ácido nucleico GRP representado pela SEQ ID N°: 396 sob o controle do promotor GOS2 mostrando um aumento de mais do que 5 % para o peso total das sementes, número de sementes enchidas, e índice de colheita, quando o desenvolvimento sob as condições de não tensão.

VII. Proteína semelhante 11 de ligação do promotor Squamosa (SPL11)

Exemplo 62: A identificação das sequências relacionadas à sequência de ácido nucleico SPL11 usados nos métodos da invenção

As sequências (cDNA eSTs de comprimento total ou genômico) relacionadas a uma sequência do ácido nucleico usado nos métodos da presente invenção onde identificados entre aqueles mantidos no banco de dados Entrez Nucleotides at the National Center for Biotechnology Information (NCB!). As sequências também buscam serem realizadas em outro banco de dados públicos e privados. As ferramentas de busca da sequência foram usadas, tal como a ferramenta de alinhamento local básica

(BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa é usado para observar as regiões da similaridade local entre as sequências em comparação ácido nucleico ou sequências de polipeptídeos de um banco de dados da sequência e pelo cálculo da significância estatística de pontos. Por exemplo, o polipeptídeo codificado pelo ácido nucleico usado na presente invenção foi usado pelo algorítmo TBLASTN, com ajustes default e o filtro para ignorar baixa complexidade sequência apresentada. A produção da análise foi examinada pela comparação em pares, e classificada de acordo com as contagens de probabilidade (valor E), onde a contagem reflete a probabilidade que um alinhamento particular ocorra pela possibilidade (o valor E inferior, quanto mais significante o acerto). Além disso aos valores E, as comparações também foram contadas pela identidade de porcentagem. A identidade da porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácido nucleico comparado (ou polipeptídeo) em um comprimento particular. Em alguns exemplos, os parâmetros default podem ser ajustados para modificar a estringência da busca. Por exemplo o valor E pode ser aumentado para mostrar menores pontos estringentes. Desta maneira, os pontos curtos próximos exatos podem ser identificados.

Tabela G1 fornece uma lista das sequências de ácidos nucleicos relacionadas à uma sequência do ácido nucleico usado nos métodos da presente invenção. Os polipeptídeos com uma acessão estende-se por um ponto e um dígito que representa as variantes de união.

Tabela G1: Exemplos de ácidos nucleicos SPL11 e polipeptídeos.

Nome	**nome alternativo**	**origem da planta**	**ácido nucleico da SEQ ID N°:**	**proteína da SEQ ID N°:**
Arath_SPL11a	At1g27360	*Arabidopsis thaliana*	427	428
Arath_SPL11aa	NM_202191 (splicing variant)	*Arabidopsis thaliana*	429	428
Arath_SPL11aaa	NM_001084135	*Arabidopsis thaliana*	430	428

Arath_SPL11aaaa	SEQID NO: 1 Variant MiR156 insensitive	Synthetic	431	428
Arath_SPL11b	At1g27370	*Arabidopsis thaliana*	432	433
Arath_SPL11c	At5g43270	*Arabidopsis thaliana*	434	435
Orysa_SPL11 a	LOC_Os02g04680	*Oryza sativa*	436	437
Orysa_SPL11 b	LOC_Os06g49010	*Oryza sativa*	438	439
Popth_SPL11a	Populus SPL11a	*Populus Species*	440	441
Popth_SPL11 b	Populus SPL11 b	*Populus Species*	442	443
Popth_SPL11c	Populus SPL11c	*Populus Species*	444	445
Brara_SPL11a	Br_AC189413	*Brassica rapa*	446	447
Zeama_SPL11 a	ZM_60752693	*Zea mays*	448	449
Zeama_SPL11 b	ZM_SPL11 b	*Zea mays*	450	451
Triae SPL11a	Ta SPL11a	*Triticum aestivum*	452	453
Vitvi SPL11a	Vv SPL11a	*Vitis vinifera*	454	455

Exemplo 63: Alinhamento das sequências de polipeptídeos SPL11

O alinhamento das sequências de polipeptídeos foi realizado usando o programa AlignX a partir do vetor NTI (Invitrogen) que é baseado no algorítmo Clustal W popular do alinhamento progressivo (Thompson et al. *(1997)* Nucleic Acids Res *25:4876-4882;* Chenna et al. *(2003).* Nucleic Acids Res 31:3497-3500). Os valores default são pela penalidade aberta de fenda de 10, pela penalidade de extensão de fenda de 0,1 e a matriz do peso selecionado é Blosum *62* (se os polipeptídeos são alinhados). Uma edição menor do manual foi feita ainda para otimizar o alinhamento. Uma sequência de consenso que compreende altamente conservados de aminoácidos representativos em cada posição é dada; o vazio é dado entre os aminoácidos na sequência de consenso representam qualquer aminoácido. A alta conservação da sequência entre os polipeptídeos SPL11 é mais observada entre o domínio SBP dos polipeptídeos. A posição de outros motivos da sequência conservada fora do domínio SBP neste representado pelo Motivo 1 a Motivo 4 (SEQ ID N°: 466 a SEQ ID N°: 472) é indicado na sequência de consenso.

Figura 28 fornece um alinhamento de polipeptídeos representativos SPL11.

Exemplo 64: Cálculo da identidade global da porcentagem entre as sequências

de polipeptídeos SPL11 úteis útil na realização de métodos da invenção

As porcentagens globais de similaridade e identidade entre as sequências de polipeptídeos de comprimento total úteis na realização de métodos da invenção foram determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Ferramenta de alinhamento global de matriz) (BMC Bioinformatics. 2003 4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando sequências de DNA ou proteínas. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). O software MatGAT gera as matrizes de similaridade/identidade para sequência de proteínas ou DNA sem necessariamento o pré-alinhamento dos dados. O programa realiza uma série de alinhamento em pares usando o algorítmo de alinhamento global Myers and Miller (com uma penalidade de abertura de fenda de 12, e a penalidade de extensão de fenda de 2), calcula a similaridade e identidade usando por exemplo Blosum 62 (para os polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade da sequência é mostrada na metada da parte inferior da linha de divisão e a identidade de sequência é mostrada na metada da parte superior da linha de divisão diagonal.

Parâmetros usados na comparação foram:

Matriz de contagem: Blosum62

Primeira fenda: 12

Extensão da fenda: 2

Os resultados da análise do software são mostrados na Tabela G2 pela similiaridade e identidade global no comprimento total de uma sequência de polipeptídeos e a Tabela G3 para a similaridade no domínio SBP. A identidade da porcentagem é dada acima em diagonal em negrito e a porcetagem da similaridade é dada abaixo em diagonal (face normal).

A identidade da porcentagem entre as sequências de polipeptídeos SPL11 úteis na realização de métodos da invenção podem ser

tão baixo quanto 20 % da identidade de aminoácido comparado a SEQ ID N°: 428. Tipicamente, polipeptídeos SPL11 tem uma identidade de sequência global (entre uma sequência do polipeptídeo inteiro) na faixa de 40 %.

A identidade da porcentagem entre os domínios SBP compreendidos nas sequências de polipeptídeos SPL11 úteis na realização de métodos da invenção podem ser tão baixo quanto 30 % da identidade de aminoácido comparado a SEQ ID N°: 456 (domínio SBP compreendido em Arath_SPL11 ou na SEQ ID N°: 428). Os domínios SBP compreendidos nos polipeptídeos SPL11 dados na Tabela G1 e tem uma identidade de sequência na faixa de 31 a 93,6 %.

Tabela G2: Resultados MatGAT para a similaridade e identidade global no comprimento total de uma sequência de polipeptídeos.

nome da sequência	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Arath SPL11 a		75,3	37,2	31,2	30,9	38,4	34,3	29,9	31	27,5	30,7	31,4	31,5
2. Arath SPL11 b	82,3		36,6	33,2	32,1	37,9	36,4	29,7	31,3	29,5	30,7	31,2	33,2
3. Arath SPL11c	52,5	52		33,6	34,9	40,4	38,6	33,7	48,5	32,9	31,8	32,4	34,4
4. Orysa_SPL11 a	46,5	44,6	46,7		54,3	36,6	38,9	38,1	25,6	46,9	49,2	66,2	34,5
5. Orysa_SPL11 b	43,8	43,4	47,8	68,2		35,4	36,7	38,4	25,2	44,3	55,1	55,2	32,6
6. Popth_SPL11a	54,5	52	54,7	49	49,1		66,5	35,4	36,4	30	32,3	35,4	44,1
7. Popth_SPL11b	46,7	48,6	52,3	55,9	53,9	71,2		37,7	32,7	30,7	34,7	37	42,5
8. Popth_SPL11c	44,7	44,7	50,4	55,9	55,6	48,9	55,7		25,8	33,3	34,2	35,7	37,6
9. Brara_SPL11 a	41,7	42,4	55,6	34,1	34,3	46,6	40	35,2		22,2	24,1	24,8	29,4
10. Zeama_SPL11a	42,1	42,5	44,6	61,4	58,3	42,7	45,9	51	31,5		41,6	49,4	28,6
11. Zeama_SPL11 b	44,2	44,4	46,7	60,6	66,7	47,6	50,5	50,4	35,1	53,1		49,5	32,1
12. Triae SPL11 a	44	44	46,3	77	69,5	48	52	53,9	33,2	62,2	61,1		32,8
13. Vitvi SPL11a	44,6	43,5	47,2	50,2	50,4	53,8	55,8	55,8	36,7	46,4	46,4	50,2	

Tabela G3: Resultados MatGAT para a similaridade e identidade no domínio SBP das sequências de polipeptídeos.

nome da seqüência	1	2	3	4	5	6	7	8	9	10	11	12
1. Arath_SPL11 a		93,7	77,2	73,4	75,9	58,3	62,2	49,5	46,6	73,4	73,4	68,2
2. Arath_SPL11 b	96,2		79,7	73,4	73,4	59,3	63,3	49,5	48,9	75,9	77,2	71,8
3. Arath_SPL11c	86,1	88,6		81	77,2	63	66,3	54,2	59,1	81	78,5	74,1
4. Orysa_SPL1 1 a	84,8	86,1	89,9		84,8	61,5	66,3	53,3	50	88,6	83,5	80
5. Orysa_SPL1 1 b	88,6	91,1	91,1	91,1		59,3	64,3	50,5	47,7	87,3	83,5	78,8
6. Popth_SPL11a	64,8	67,6	68,5	65,7	66,7		60,6	52	45,4	62,4	57,4	60,2
7. Popth_SPL11 b	69,4	72,4	72,4	70,4	72,4	71,3		45,2	41,3	66,3	62,2	61,5
8. Popth_SPL11c	63,3	64,3	65,3	64,3	66,3	66,7	65,3		31	53,3	50,5	56,1
9. Brara_SPL11 a	62	64,6	68,4	63,3	64,6	56,5	55,1	43,9		51,1	51,2	46,8
10. Zeama_SPL11 a	86,1	88,6	91,1	93,7	96,2	66,7	70,4	67,3	63,3		88,6	84,7
11. Zeama_SPL11 b	88,6	91,1	89,9	88,6	93,7	63,9	69,4	64,3	64,6	92,4		90,6
12. Triae SPL11 a	82,4	84,7	84,7	84,7	88,2	67,6	72,4	71,4	60	88,2	90,6	
13. Vitvi SPL11 a	90,1	93,8	90,1	87,7	90,1	70,4	75,5	67,3	64,2	88,9	87,7	87,1

Exemplo 65: A identificação dos domínios compreendidos nas sequências de polipeptídeos úteis na realização de métodos da invenção

O recurso integrado das famílias de proteínas, os bancos de dados de locais e domínios (InterPro) é uma interface integrada para os bancos de dados assinados comumente usados para as buscas com base no texto e sequência. O banco de dado InterPro combina estes bancos de dados, que usam as metodologias diferentes e variam os graus de informação biológica sobre as proteínas bem caracterizadas para derivar os sinais da proteína. A colaboração dos bancos de dados incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. O Pfam é uma ampla coleção de alinhamentos de sequência múltipla e modelos Markov secretos que cobrem muitos dos domínios de proteína comuns e famílias. O Pfam é hospedado no servidor Sanger Institute in the United Kingdom. O Interpro é hospedado no European Bioinformatics Institute in the United Kingdom.

Tabela G4 mostra os ajustamentos (corte de União, corte de confiança e corte de ruído) como descrito nos bancos de dados Pfam que foram usados para construir os HMMs_fs para os domínios diferentes SBP.

Tabela G4: Informação de HMMs_construção em bancos de dados do domínio Pfam Versão 21.0

informação de construção HMMER		
	Pfam_ls [HMM carregado]	Pfam_fs [HMM carregado]
corte União	25.0 25.0;	25.0 25.0
corte de confiança	41.4 41.4;	26.2 54.1
corte de ruído	20.8 20.8;	18.0 11.2
método de construção de HMM	hmmbuild -F HMM ls SEED	hmmbuild -f -F HMM_fs SEED
	hmmcalibrate --seed 0 HMM ls	hmmcalibrate --seed 0 HMM_fs

Os resultados da avaliação Pfam para os polipeptídeos representativos SPL11 de origem vegetal são apresentados na Tabela G5. As coordenadas do aminoácido para o domínio SBP na sequência de referência é indicado na coluna correspondente. O valor E do alinhamento também é dado.

Tabela G5: Resultados da avaliação Pfam para o domínio SBP (Pfam referência PF03110) de polipeptídeos representativos SPL11 como realizados

em Pfam Versão 21.0.

nome da sequência	coordenadas de aminoácido do domínio SBP	valor E do alinhamento
Arath SPL11 a	174-252	1.10E-50
Arath SPL11 b	175-253	1.61E-47
Arath SPL11c	168-246	1.90E-52
Orysa_SPL11 a	181-259	5.10E-50
Orysa_SPL11 b	179-257	3.90E-47
Popth_SPL11 a	170-277	170-277
Popth_SPL11 b	180-277	1.50E-41
Popth_SPL11 c	171-249	7.80E-52
Brara SPL11 a	180-279	5.20E-44
Zeama SPL11a	161-239	2.20E-51
Zeama SPL11 b	159-237	1.10E-48
Triae SPL11 a	184-262	2.90E-52
Vitvi SPL11a	171-249	7.80E-52

Exemplo 66: Clonagem da sequência do ácido nucleico usado nos métodos da invenção

A sequência do ácido nucleico usado nos métodos da invenção foi amplificado por PCR usando como modelo uma feitas comumente na biblioteca de cDNA de mudas *Arabidopsis thaliana* (em pCMV Sport 6.0; Invitrogen, Paisley, UK). O PCR foi realizado usando a polimerase de DNA Hifi Taq em condições padrões, usando 200 ng do modelo em uma mistura de PCR a 50 µl. Um iniciador a montante e a jusante como representado pela SEQ ID N°: 473 e SEQ ID N°: 474 respectivamente foram usados para amplificar por PCR (Reação de cadeia de polimerase) a região codificadora de um ácido nucleico SPL11 como representado pela SEQ ID N°: 427. Os iniciadores incluem os locais AttB para a recombinação Gateway para facilitar a clonagem do fragmento de DNA de PCR amplificado no vetor de clonagem Gateway.

O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante o qual o fragmento PCR recombina *in vivo* com o plasmídio pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada", pSPL11. O plasmídio

pDONR201 foi adquirido de Invitrogen, como partes da tecnologia Gateway®.

O clone de entrada que compreende SEQ ID N°: 427 foi então usado em uma reação LR com um vetor de destino usado para a transformação de *Oryza sativa*. Este vetor contido como os elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; um cassete de expressão marcador avaliável; e um cassete Gateway pretendido pela recombinação LR *in vivo* com uma sequência do ácido nucleico de interesse já clonado no clone de entrada. Um promotor GOS2 de arroz (SEQ ID N°: 476) para a expressão constitutiva foi localizada a montante deste cassete Gateway no vetor pGOS2::SPL11. A sequência codificadora SPL11 também foi clonada em um vetor de expressão que compreende um promotor WSI18 de arroz (SEQ ID N°: 477) para a expressão específica de semente (ABA induzível) (pWSI18::SPL11).

Após a etapa de recombinação LR, o vetor de expressão resultante pGOS2::SPL11 e pWSI18::SPL11 (Figura 31) foram transformados em cepa de *Agrobacterium* LBA4044 de acordo com os métodos bem conhecidos na técnica.

Exemplo 67: Transformação da planta

A transformação das plantas foi realizada de acordo com os procedimentos resumidos no Exemplo 7

Exemplo 68: Procedimento de avaliação fenotípica

68.1 Apresentação da avaliação

Aproximadamente 35 transformantes de arroz TO independentes foram gerados. Os primeiros transformantes foram transferidos a partir de uma câmera de cultura de tecido de uma estufa para o desenvolvimento e coletado da semente T1. Seis eventos, de que uma progênie T1 segregada 3:1 para a presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 mudas T1

contendo o transgene (hetero e homozigotos) e aproximadamente 10 mudas T1 perderam o transgene (nulizigotos) foram selecionados pela expressão marcadora visual do monitoramento. As plantas transgênicas e os nulizigotos correspondentes foram desenvolvidos lado a lado em posições aleatórias. As condições da estufa foram de dias curtos (12 horas de luz), 28° C na luz e 22° C no escuro, e uma umidade relativa de 70 %. O desenvolvimento das plantas sob as condições de não tensão são fornecidos com água em intervalos regulares para garantir que a água e nutrientes não são limitantes para satisfazer a necessidade das plantas para completar o crescimento e desenvolvimento.

Quatro eventos T1 ainda foram avaliados na geração T2 seguindo o mesmo procedimento de avaliação como para a geração T1 mas com mais indivíduos por evento. A partir do estágio de semeadura até o estágio de maturidade das plantas foram passados diversas vezes através de um gabinete de imagem digital. Em cada período das imagens digitais (2048x1536 píxeis, 16 milhões de cores) foram tiradas de cada planta em pelo menos 6 ângulos diferentes.

Avaliação da aridez

As plantas das sementes T2 foram desenvolvidas no solo em potes sob condições normais até estes aproximarem do estágio superior. Estes foram então transferidos à seção "seca" onde a irrigação foi retida. As sondas de umidade foram inseridas em potes escolhidos aleatoriamente para monitorar o teor de água no solo (SWC). Quando o SWC está abaixo de certos limiares, as plantas foram automaticamente novamente submetidas à água continuamente até um nível normal foi atingido novamente. As plantas foram então re-transferidas novamente a condições normais. O resto do cultivo (amadurecimento da planta, coleta da semente) foi o mesmo para as plantas não desenvolvidas sob as condições de tensão abiótica. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais.

Avaliação da eficiência do uso de nitrogênio

As plantas de arroz das sementes T2 são desenvolvidas no solo em potes sob condições normais exceto para a solução do nutriente. Os potes são submetidos à água a partir da transplantação do amadurecimento com uma solução de nutriente específico contendo teor de nitrogênio N reduzido (N), usualmente entre 7 a 8 vezes menor. O resto do cultivo (amadurecimento da planta, coleta da semente) é o mesmo como para as plantas não desenvolvidas sob a tensão abiótica. Os parâmetros de produção e desenvolvimento são registrados como detalhado para o desenvolvimento sob as condições normais.

68.2 Análise estatística: Teste F

Dois fatores ANOVA (análise de variantes) foram usados como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformados com o gene da presente invenção. O teste F foi realizado para checar por um efeito do gene em todos os eventos de transformação e para verificar por um total do efeito do gene, também conhecido como um efeito do gene global. O limiar para a significância por um efeito verdadeiro do gene global foi apresentado em uma 5 % do nível de probabilidade para o teste F. Um valor significante do teste F por um efeito do gene, significando que este não é apenas uma mera presença ou posição do gene que está causando as diferenças no fenotipo.

Quando dois experimentos com os eventos de sobreposição foram realizados, uma análise combinada foi realizada. Este é útil para checar a consistência dos efeitos nos dois experimentos, e se este é o caso, para acumular a evidência de ambos experimentos a fim de aumentar a confidência da conclusão. O método usado foi um método modelo misturado que leva em conta a estrutura multinível dos dados (isto é experimento - evento - segregantes). Os valores P obtidos em comparação ao teste de razão de probabilidade para distribuições quadradas chi.

68.3 Parâmetros medidos

Medição do parâmetro relacionado à biomassa

A partir do estágio de semeadura até o estágio de maturidade das plantas foram passados diversas vezes através de um gabinete de imagem digital. Em cada período das imagens digitais (2048x1536 píxeis, 16 milhões de cores) foram tiradas de cada planta em pelo menos 6 ângulos diferentes.

Uma área da planta acima do solo (ou biomassa da folha) foi determinado pela contagem do número total de píxeis nas imagens digitais a partir das partes de plantas acima do solo discriminadas a partir do fundamento. Este valor foi medido pelas figuras tiradas no mesmo tempo a partir de ângulos diferentes e foi convertido a um valor de superfície física expressada em quadrado mm pela calibração. Os experimentos mostram que à área da planta acima do solo medida desta maneira correlaciona com a biomassa das partes das plantas trituradas acima. À área triturada acima é a área medida no período em que a planta tem atingido esta biomassa máxima da folha. O vigor precoce é à área das planta (mudas) acima do solo três semanas pós-germinação. O aumento na biomassa da raiz é expressado como um aumento no total da biomassa da raiz (medido como biomassa máxima de raízes observadas durante a vida toda de uma planta); ou como um aumento na raiz/índice de broto (medido como a razão entre massa de raiz e massa de broto no período de desenvolvimento ativo da raiz e broto).

O vigor emergente foi determinado pela contagem do número total de píxeis a partir das partes de plantas acima do solo discriminadas a partir do fundamento. Este valor foi medido pelas figuras tiradas no mesmo tempo de ângulos diferentes e foi convertido a um valor de superfície física expressada em quadrado mm pela calibração. Os resultados descritos abaixo são para as plantas com três semanas pós-germinação.

Medições dos parâmetros relacionados à semente

As panículas primárias maduras foram coletadas, contadas, ensacadas, rotuladas com código de barra e então secadas por três dias em um

forno a 37° C. As panículas foram então limiadas e todas as sementes foram coletadas e contadas. As cascas enchidas foram separadas a partir de um vazio usando um dispositivo de sopro de ar. As cascas vazias foram descartadas e a fração remanescente foi contada novamente. As cascas enchidas foram pesadas em uma balança analítica. O número de sementes enchidas foi determinado pela contagem do número de cascas enchidas que permanece depois da etapa de separação. A produção total de semente foi medida pela pesagem total das cascas enchidas coletadas a partir da planta. O número total de semente por planta foi medido pela contagem do número de cascas coletadas a partir da planta. Peso de núcleo milhar (TKW) é extrapolado do número de sementes enchidas contadas e seu peso total. O índice de colheita (HI) na presente invenção é definida como a razão entre a produção total de semente e à área triturada acima (mm^2), multiplicada por um fator 10^6. O número total de flores por panícula como definido na presente invenção é a razão entre o número total de sementes e o número de panículas primárias maduras. A taxa de abastecimento da semente como definido na presente invenção é a proporção (expressado como uma %) do número de sementes enchidas no número total de sementes (ou florzinhas).

Exemplo 69: Resultados da avaliação fenotípica das plantas transgênicas

Os resultados da avaliação das plantas transgênicas de arroz que expressam um ácido nucleico SPL11 correspondente à construção de pGOS2::SPL11 e pWSI18::SPL11 se sob a tensão da aridez ou condições de não tensão são dados neste. Um aumento em pelo menos 5 % foi observado pela um ou mais dos seguintes parâmetros de vigor emergente (mudas com vigor precoce), produção total da semente (peso da semente), a taxa de abastecimento da semente (taxa de enchimento da semente), o número de sementes enchidas, o número de flores (sementes) por panícula e o índice de colheita, e em pelo menos 3 % por mil grãos em peso nas plantas transgênicas (1000 sementes) quando comparados ao controle de plantas nulizigotos em pelo menos uma das condições de desenvolvimento.

MSCNGCRVLRKGCSENCILRPCIQWIETADAQGHATVFVAKFFGRAGLMSFISAVPDS

1 2

QRPALFQSLLYEACGRTVNPVNGAIGMLWTGNWNICQAAVETVLRGGSLRPIPELLTH

3

GGGFAGFPSPTSEEASEICTEMLNLQQNDSTDRNIYHHSRFSSSRSRSTMDSSSPTKRK

RLSSEDQPSSELDLSLIPNFPIKQATPSSTRRRSVTPSMNSEDSGTTTTTTAFCDKGDVY

GNGGGETTKLLNLFV

FIGURA 1

CLUSTAL W (1.83) **Alinamento de sequência múltiplo**

```
Os-Os01g03890          ----------------------------------------MRMSCNGCRVLRKGCSDN
St-TC162239            ----------------------------------------MRMSCNGCRVLRKGCSEN
Ta-TC235711            ----------------------------------------MRLSCNGCRVLRKGCSED
Hv-TC147808            ----------------------------------------MRLSCNGCRVLRKGCSED
Os-Os01g32770          ----------------------------------------MRMSCNGCRVLRKGCSEG
At-LBD42-Q9CA30        ----------------------------------------MRISCNGCRVLRKGCNQD
Pt-123.86              ----------------------------------------MRMSCNGCRVLRKGCSEN
At-LBD40-Q9ZW96        ----------------------------------------MRMSCNGCRVLRKGCSEN
Bn-AAP37970            ----------------------------------------MRMSCNGCRVLRKGCSDD
At-LBD41-Q9M886        ----------------------------------------MRMSCNGCRVLRKGCSED
Mt-TC107091            ----------------------------------------MRMSCNGCRVLRKGCSEN
St-TC133081            ----------------------------------------MRMSCNGCRVLRKGCSKS
Gm-TC220806            ----------------------------------------MRMSCNGCRVLRKGCSED
Mt-ABE78739            ----------------------------------------MRMSCNGCRVLRKGCSED
Vv-TC55931             ----------------------------------------MRMSCNGCRVLRKGCSDD
Aq-TC18329             ----------------------------------------MRMSCNGCRVLRKGCSEN
Pt-X180                ----------------------------------------MSCNGCRVLRKGCNDN
Pt-XII.481             MPFSALSFSSLPHRIILLLIGFLNSSPSLSRHFSISFRTEYTMRMSCNGCRILRKGCGDN
Pta-TC67269            ----------------------------------------MRMSCNGCRVLRKGCSET
Pta-TC69225            ----------------------------------------MRVSCNGCRVLRKGCSDS
Os-Os03g33090          ------------------------------------------MSCNGCRVLRKGCSDG
Sb-TC104758            ------------------------------------------MSCNGCRVLRKGCSEA
Os-Os07g40000          ------------------------------------------MSCNGCRVLRKGCSEG
Hv-TC147776            ------------------------------------------MSCNGCRVLRKGCNDA
Os-Os03g41330          ------------------------------------------MSCNGCRVLRKGCSDA
Ta-TC270332            ------------------------------------------MSCNGCRVLRKGCSDA
St-TC137193            ------------------------------------------MSCNGCRVLRKGCSDN
St-TC133385            ------------------------------------------MSCNGCRVLRKGCSDN
Le-TC178827            ------------------------------------------MSCNGCRVLRKGCSDN
Nb-TC9404              ------------------------------------------MSCNGCRVLRKGCSEN
Gm-TC216138            ------------------------------------------MSCNGCRVLRKGCSES
Lj-TC15459             ------------------------------------------MSCNGCRVLRKGCSES
Gm-TC227562            ------------------------------------------MSCNGCRVLRKGCSES
Ck-LOB-DQ787782        ------------------------------------------MSCNGCRVLRKGCSES
Mt-ABE82505            ------------------------------------------MSCNGCRVLRKGCSES
Pt-III.589             ------------------------------------------MSCNGCRVLRKGCSEN
Pt-V.543               ------------------------------------------MSCNGCRVLRKGCSEN
Vv-TC60668             ------------------------------------------MSCNGCRVLRKGCSES
Gh-TC30552             ------------------------------------------MSCNGCRVLRKGCSES
At-LBD37-Q9FN11        ------------------------------------------MSCNGCRVLRKGCSEN
At-LBD38-Q9SN23        ------------------------------------------MSCNGCRVLRKGCSEN
At-LBD39-Q9SZE8        ------------------------------------------MSCNGCRVLRKGCSET
Aq-TC18561             ------------------------------------------MSCNGCRVLRKGCSES
Pta-TC59178            ---------------MTATTGAGNSGRDLSVFIQRKQEADGRKRMSCNGCRVLRKGCSDT
St-TC140088            ------------------------------------------MSCNGCRILRKGCNEN
Pa-TC14656             -------------------------MSATDKNITERKEQQRKMSCNGCRVLRKGCSDN
Pta-TC67974            -------------------------MSATENHITERKEQERKMSCNGCRVLRKGCSEN
Pt-XIV.94              ------------------------------------------MSCNGCRVLRKGCSET
Pt-II105               ------------------------------------------MSCNGCRVLRKGCSET
                                                                  :******:******..
```

FIGURA 2

```
Os-Os01g03890   CAIRPCLQWIRSPDAQANATVFLAKFYGRAGLINLITAG-PEHVRPAIFRSLLYEACGRM
St-TC162239     CSLRPSLQWIKSPDSQANATVFLAKFYGRAGLINLINAG-PDHLRPAIFRSLLYEACGRI
Ta-TC235711     CSIRPCLQWIKSPEAQANATVFLAKFYGRAGLMNLINAGTDDSLRPGIFRSLLYEACGRI
Hv-TC147808     CSIRPCLQWIKSPEAQANATVFLAKFYGRAGLMNLINAGTDDSLRPGIFRSLLYEACGRI
Os-Os01g32770   CTIRPCLQWIKTPEAQANATVFLAKFYGRAGLLNLLAAG-PDHLRPAVFRSLLYEACGRI
At-LBD42-Q9CA30 CTIRPCLQWIKSADSQANATLFLAKFYGRAGLLNLIESG-PDHLRPAIFRSLLYEACGRI
Pt-123.86       CSIRPCLQWIKSSESQANATVFLAKFYGRAGLMNLINAG-PEHLRPAIFRSLLYEACGRI
At-LBD40-Q9ZW96 CSIRPCLQWIKSAESQANATVFLAKFYGRAGLMNLLNTG-PDHLRPAIFRSLLYEACGRI
Bn-AAP37970     CSIRPCLGWIKSPEAQANATVFLAKFYGRAGLMNLINAG-PDHLRPGIFRSLLHEACGRI
At-LBD41-Q9M886 CSIRPCLAWIKSPEAQANATVFLAKFYGRAGLMNLINAG-PNHLRPGIFRSLLHEACGRI
Mt-TC107091     CSIRPCLQWIKSPESQANATVFLAKFYGRAGLMNLVNAG-PEHLRPAIFRSLLYEACGRI
St-TC133081     CSIRPCLQWIKTPDSQSNATVFLAKFYGRAGLMNLINAG-PDHLRPAIFRSLLYEACGRI
Gm-TC220806     CSIRPCLQWIKNPESQANATVFLAKFYGRAGLMNLINAG-PENLRPAIFRSLLYEACGRI
Mt-ABE78739     CSIRPCLEWIKCPQSQANATLFLAKFYGRAGLINLINSG-SENLRPAIFRSLLYEACGRI
Vv-TC55931      CSLRPCLEWIKTPESQANATVFLAKFYGRAGLMNLLNAG-PEHLRPAIFRSLLYEACGRI
Aq-TC18329      CSIRPCLQWIKTPESQANATVFLAKFYGRAGLMNLINSG-PEHLRPAIFRSLLYEACGRI
Pt-X180         CTIRPCLQWIKSTDSQANATLFLAKFYGRAGLINLIEAA-PQRLRPAIFSSLLYEACGRI
Pt-XII.481      CSIKPCLQWIETPDSQANATLFLAKFYGRAGLMNLINAC-PQHLRPDTFKSLLYEACGRI
Pta-TC67269     CIIRPCLQWIKNEESQANATVFLAKFYGRAGLMNLISAG-PEHLRPAIFKSLLYEACGRM
Pta-TC69225     CELRPCLQWIKTSESQAQATVFLAKFYGRSGLMNLIKSG-PEFLRPTLFRSLLYEACGRI
Os-Os03g33090   CVLRPCLQWIDAADAQGHATVFVAKFFGRAGLLSFISAV-PEAQRPALFQSLLYEAAGRT
Sb-TC104758     CVLRPCLQWIDAADAQGHATVFVAKFFGRAGLLSFISAV-PDAQRPALFQSLLYEAAGRT
Os-Os07g40000   CVLRPCLQWIDGAEAQGHATVFVAKFFGRAGLMSFLTAV-PEPQRAAVFQSLLYEAAGRT
Hv-TC147776     CMLRPCLLWIEGADAQGHATIFAAKFFGRAGLMSFLTAV-PESQRPAVFQSLLYEAAGRT
Os-Os03g41330   CVLRPSIEWIDGAQPQANATVFVAKFFGRAGLVASLAAV-PLHHRPALFRSLLYEACGRT
Ta-TC270332     CVLRPSIEWIDGAQPQANATVFVAKFFGRAGLVASLAAV-PLHHRPALFRSLLYEACGLT
St-TC137193     CILRPCLQWIETPEAQGHATVFVAKFFGRAGLMSFISAV-PENQRPALFQSLLYEAAGRT
St-TC133385     CILRPCLQWIETPEAQGHATVFVAKFFGRAGLMSFISAV-PENQRPALFQSLLYEAAGRT
Le-TC178827     CILRPCLQWIETPEAQGHATVFVAKFFGRAGLMSFISAV-PENQRPALFQSLLYEAAGRT
Nb-TC9404       CILRPCLQWIETAESQGHATVFVAKFFGRAGLTSFISAV-PENQRPSLFQSLLYEAAGRT
Gm-TC216138     CILRPCLQWIDTPEAQGHATVFVAKFFGRADLMSFISNV-PEPQRPALFQSLLFEACGRT
Lj-TC15459      CILRPCLQWIDTAEAQGHATVFVAKFFGRADLMSFISNV-PQPQRPSLFQSLLFEACGRT
Gm-TC227562     CILRPCLQWIETAEAQGHATVFVAKFFGRAGLMSFISNV-PENQRPALFQSLLFEACGRT
Ck-LOB-DQ787782 CILRPCLQWLENPEAQGHATVFVAKFFGRAGLMSFLSNV-PETQRPALFQSLLFEACGRT
Mt-ABE82505     CILRPCLQWIETPEAQGHATVFVAKFFGRAGLMSFISNV-PEPQRPALFQSLLFEACGRT
Pt-III.589      CILRPCLQWIESPEAQGHATVFVAKFFGRAGLMSFISAV-PENQRPSLFQSLLFEACGRT
Pt-V.543        CILRPCLQWIESPEAQGHATVFVAKFFGRAGLMSFISSV-PEDQRPSLFQSLLFEACGRT
Vv-TC60668      CILRPCLQWIESAESQGHATVFVAKFFGRAGLMSFISAV-PENQRPALFRSLLYEACGRT
Gh-TC30552      CILRPCLQWIESPEAQGHATVFVAKFFGRAGLMSFISAV-PESQRPALFQSLLFEACGRT
At-LBD37-Q9FN11 CILRPCIQWIETADAQGHATVFVAKFFGRAGLMSFISAV-PDSQRPALFQSLLYEACGRT
At-LBD38-Q9SN23 CILRPCIQWIESPEAQGHATVFVAKFFGRAGLMSFISAV-PESQCPALFQSLLYEACGRT
At-LBD39-Q9SZE8 CILRPCLQWIESAESQGHATVFVAKFFGRAGLMSFISSV-PELQRPALFQSLLFEACGRT
Aq-TC18561      CILRPCLQWIESPEAQGHATVFIAKFFGRAGLMSFISSV-PESQRPALFQSLLFEACGRT
Pta-TC59178     CILRPCLQWIESPEAQGHATVFVAKFFGRAGLMSFISAV-PDHQRPALFQSLLYEACGRT
St-TC140088     CILRQSLQGIESPHAQANATVFVAKFFGRAGLMSFLTSV-SESQRPALFQSLLFEACGRT
Pa-TC14656      CILRQSLQWIESPQSQAHATVFVAKFFGRAGLMTFIATV-SEPERPDLFKSLLYEACGRT
Pta-TC67974     CILRQSLQWIESPQSQAHATVFVAKFFGRAGLMTFIAAV-SEPERPALFKSLLYEACGRT
Pt-XIV.94       CVLRSCLHWIPTPEAQGNATLFLAKFFGRSDLISLISAV-PESQRPVLFQSLLFEACGRT
Pt-II105        CVLRSCLHWITNPEAQGNATLFLAKFFGRSDLMSLISAV-PEAQRPALFQSLLFEACGRT
                * :: .:  :   ..*.:**:* ***:**:.*   :          .  * ***.**.*
```

FIGURA 2 CONT.

```
Os-Os01g03890     LNPVYGSVGLLWSGNWQLCQSAVESVLRGMPIAQPPPS----------
St-TC162239       INPVNGSVGLMCSGNWQRCQEAVESVLKGSTIMQVPINNDDDNN----------
Ta-TC235711       VNPIYGSVGLLWSNNWQMCQAAVEAVLSGKPIVQVSS----------
Hv-TC147808       VNPIYGSVGLLWSNNWQMCQAAVEAVLSGKPIVQVSS----------
Os-Os01g32770     VNPIYGSVGLLWSGQWQACQAAVEAVLKGDPVVQVSS----------
At-LBD42-Q9CA30   VNPVDGSVGLMWSGNWAQCQAAVDAVLNGLPITHTPLP----------
Pt-123.86         VNPIYGSVGLMWSGSWQLCQAAVEAVLKGAPITPINSEAA----------
At-LBD40-Q9ZW96   VNPIYGSVGLLWSGNWHLCQAAVEAVMRGSPVTPIACDAA----------
Bn-AAP37970       VNPIYGSVGLLWSGNWQLCQSAVEAVMKGEPITEMATDAA----------
At-LBD41-Q9M886   VNPIYGSVGLLWSGNWQLCQDAVEAVMKGEPVKEIATDAA----------
Mt-TC107091       VNPIYGSVGLLWSGSWQLCQAAVEAVLKGAPITPITSEAA----------
St-TC133081       VNPIYGSVGLLWSGNWQLCQNAVEAVLKGTPITPIASEIA----------
Gm-TC220806       VNPIYGSVGLLWSGSWQLCQAAVEAVLKGEPITPITSEAA----------
Mt-ABE78739       VNPIYGSVGLLWSGSWHLCQAAVEAVLKGEPITPIDSEAD----------
Vv-TC55931        VNPIYGSVGLLWSGNWNHCQNAVDAVLGGSQIMEVPCDTA----------
Aq-TC18329        VNPIYGSVGLLWSGSWQLCQAAVEAVLKGTPIMQISSEVA----------
Pt-X180           VNPVYGSVGMLWSGNWAQCQAAVDAVLKGLPIISIPSSDVPTP----------
Pt-XII.481        VNPVSGSVGLLSTGSWQQCQAAVEAVLKGEPITQIASTDQLT----------
Pta-TC67269       VNPIFGSVGLLCSGNWQLCQAAVESILKGCPIHPILVDSSGTQM----------
Pta-TC69225       VNPIFGSVGLIWSGNWHICQAAVESVLKGYPLQPLHPLNSLEIA----------
Os-Os03g33090     INPVHGAVGLLWTGNWPLCQAAVETVLRGGAIGPLP----------
Sb-TC104758       INPVHGAVGLLGTGNWHLCQAAVDTVLRGGAIGPLP----------
Os-Os07g40000     INPVGGAVGLLSGGSWHLCQAAVDTVLRGGGIQPLP----------
Hv-TC147776       INPVGGAVGLLWAGSWHLCEAAVQTVLRGGAIRPLP----------
Os-Os03g41330     INPVSGAIGLMWTGNWDLCQAAADAVLRGDSLRALS----------
Ta-TC270332       INPVSGAIGLMWTSNWDLCQAAAEAVLRGDSLRSLS----------
St-TC137193       VNPVNGAVGLLWTGNWHVCQAAVETVLRGGALRPIS----------
St-TC133385       VNPVNGAVGLLWTGNWHVCQAAVETVLRGGALRPIS----------
Le-TC178827       VNPVNGAVGLLWTGNWHVCQAAVETVLRGGALRPIS----------
Nb-TC9404         VNPVNGAVGLLWTGNWHVCQAAVETVLRGGALRPIP----------
Gm-TC216138       VNPVNGAVGLLWTGNWHVCQAAVETVLRGGTVRPMP----------
Lj-TC15459        VNPVNGAVGLLWTGNWHVCQAAVETVLRGGTLKPLP----------
Gm-TC227562       VNPVNGAVGLLWTGNWHVCQAAVETVLRGGTLRPMP----------
Ck-LOB-DQ787782   VNPVNGAVGLLWTGNWHVCQAAVETVLQGGTLRPLP----------
Mt-ABE82505       VNPVNGAVGLLWTGNWHVCQAAVETVLRGGTLRPLP----------
Pt-III.589        VNPVNGAAGLLWTGNWHVCQAAVETVLRGGTLQPMP----------
Pt-V.543          VNPVNGAVGLLWTGNWHVCQAAVETVLRGGTLRPMP----------
Vv-TC60668        VNPVNGAVGLLGTGNWHVCQAAMETVLRGGTLRAMP----------
Gh-TC30552        VNPVNGAVGLLWNGNWHVCQAAVETVLRGGSLRSMP----------
At-LBD37-Q9FN11   VNPVNGAIGMLWTGNWNICQAAVETVLRGGSLRPIP----------
At-LBD38-Q9SN23   VNPVNGAVGLLWTGNWNVCQAAVETVLRGGSLKPIP----------
At-LBD39-Q9SZE8   VNPVNGAVGMLWTRNWHVCQAAVETVLRGGTLRPIS----------
Aq-TC18561        VNPVNGAVGLLWTGNWHICQAAVETVLRGGTLRPIN----------
Pta-TC59178       VNPVNGAVGLLWNGNWQLCQAAVETVLKGGTLRPPQSSSVIPTCSTRPSLMNSFSANFRV
St-TC140088       VNPVHGAVGLLWTGNWHVCQSAVETVLKGGVLRPLP----------
Pa-TC14656        VNPVHGAVGLMWTGKWQICQAAVQTVLSGGSLDRTPAS----------
Pta-TC67974       VNPVHGAVGLLWTGKWQICQAAVQTVLSGGSLDRTPAG----------
Pt-XIV.94         VNPVNGAVGLLWSGNWHVCQAAVESVLAGETLRPLPG----------
Pt-II105          VNPVNGVVGLLWSGNWHMCQEAVETVLSGGALQPLPG----------
                  :**: *  *::   .*  *: * :::: *  :
```

FIGURA 2 CONT.

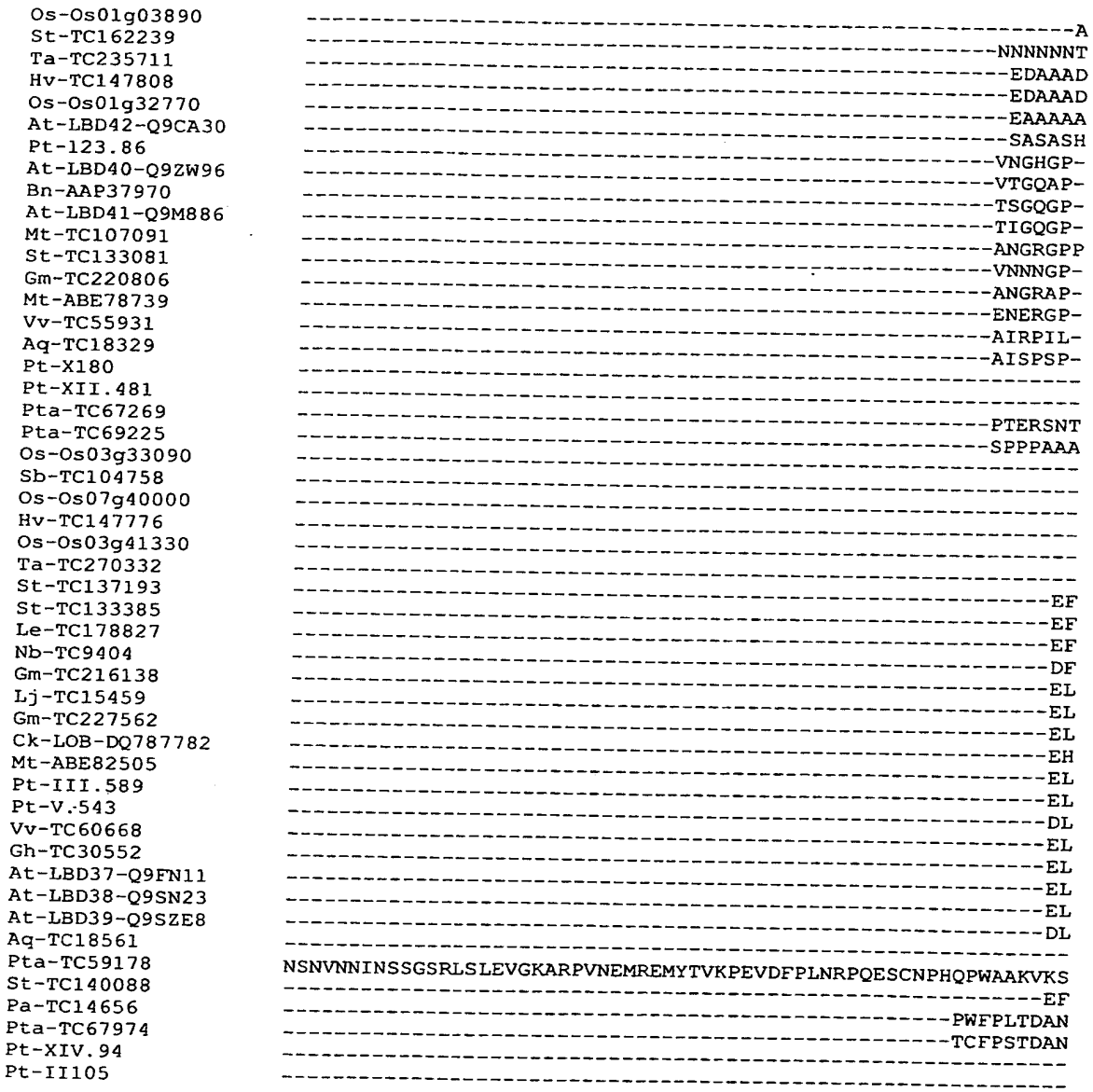


FIGURA 2 CONT.



FIGURA 2 CONT.

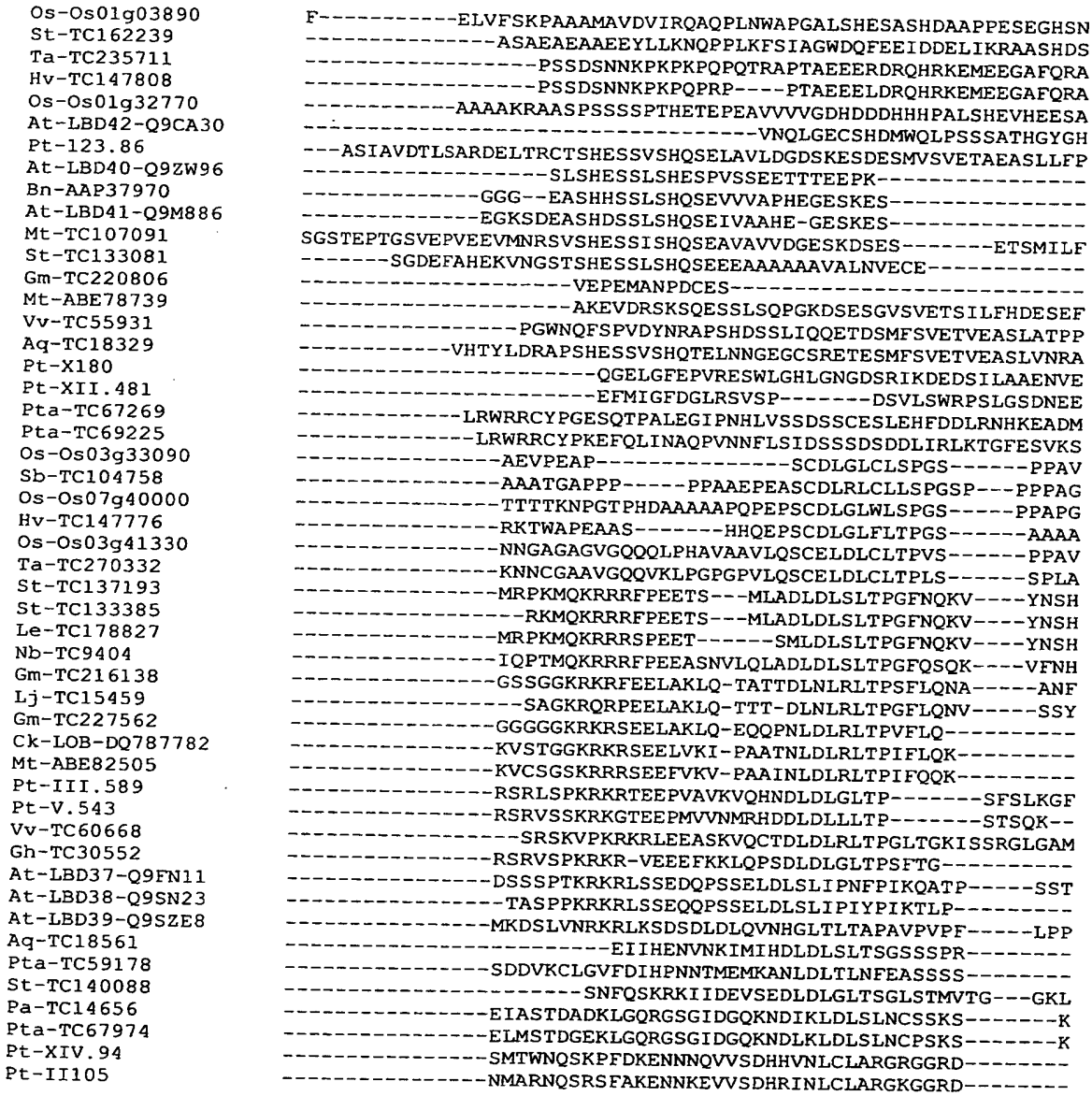


FIGURA 2 CONT.

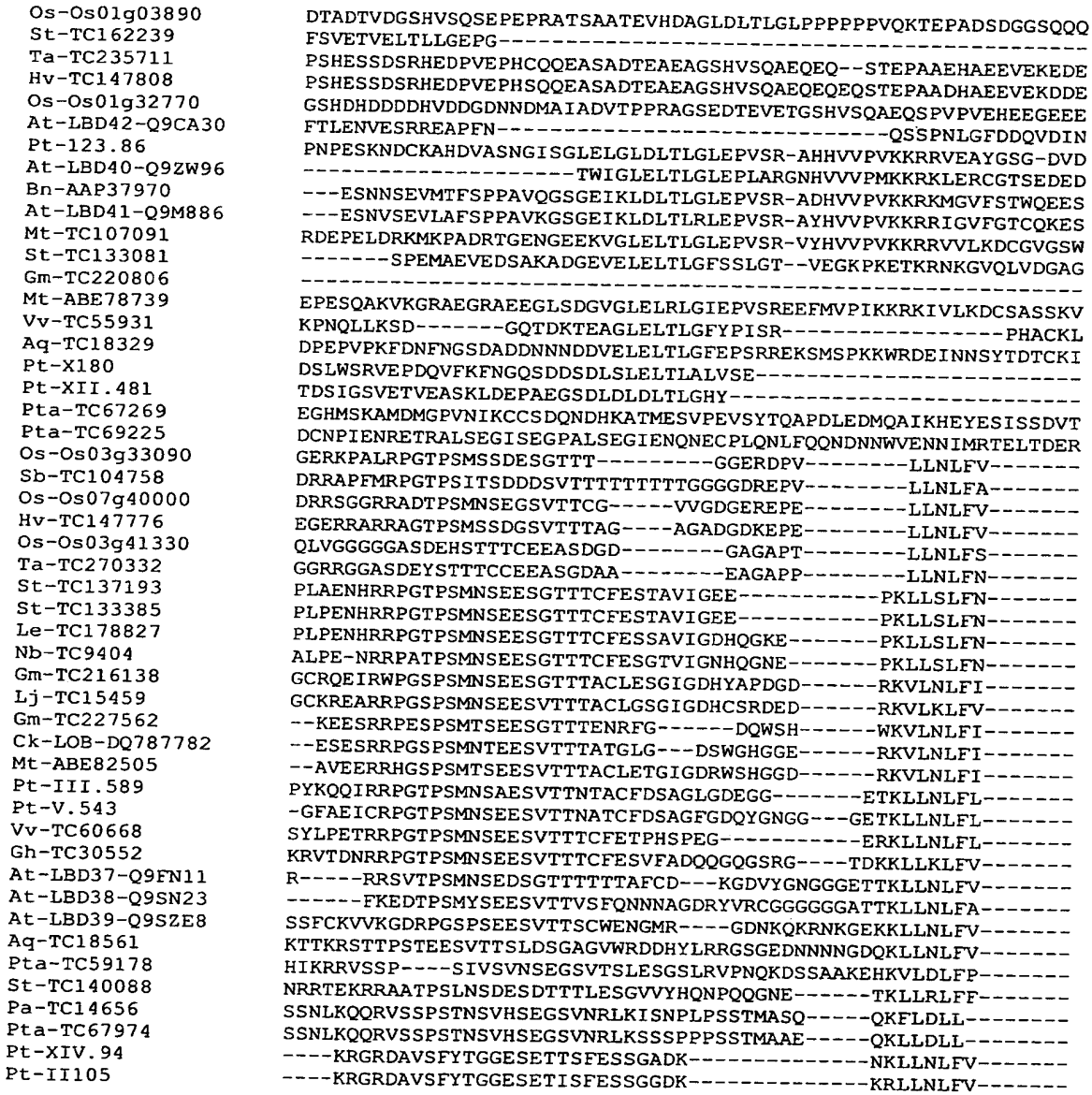


FIGURA 2 CONT.

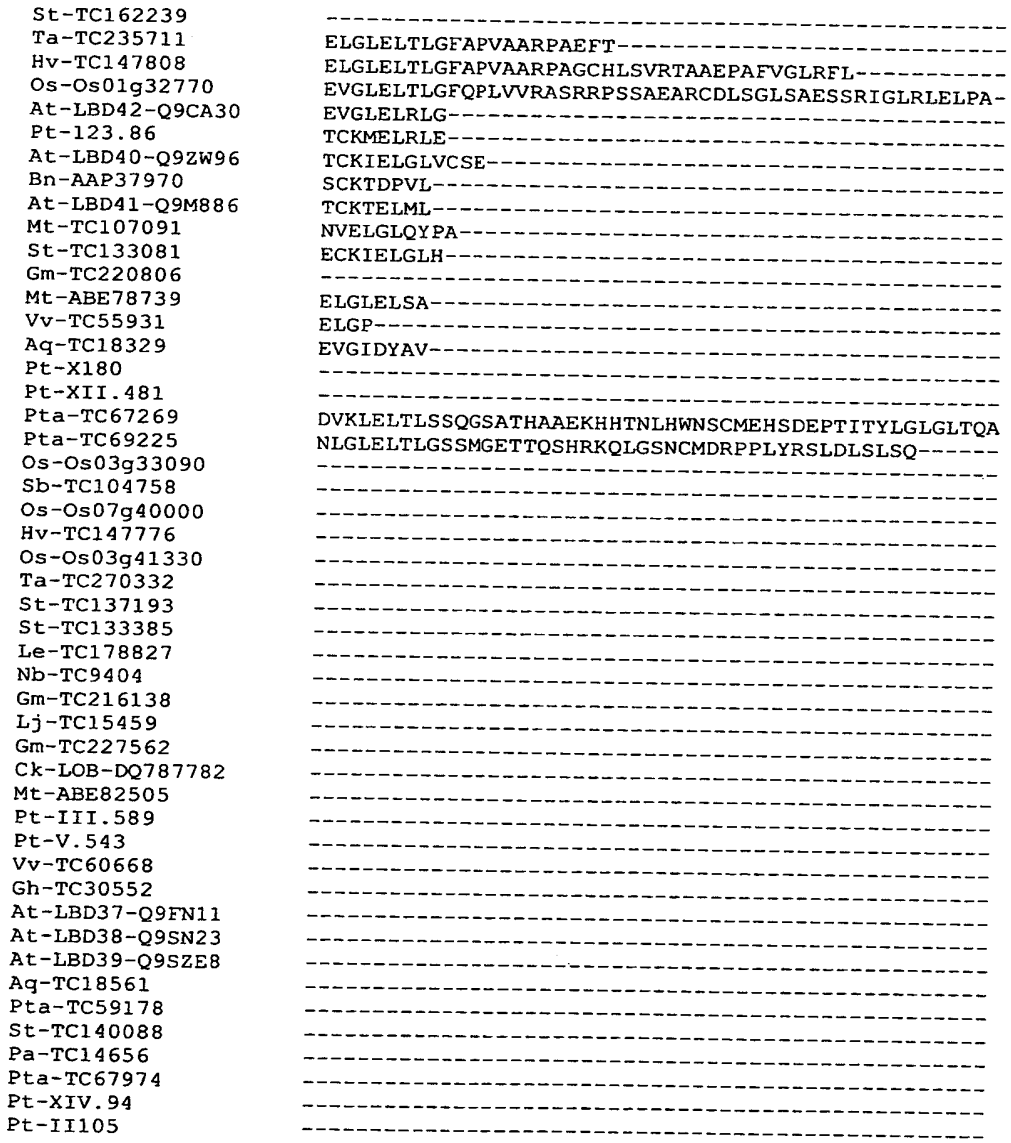


FIGURA 2 CONT.

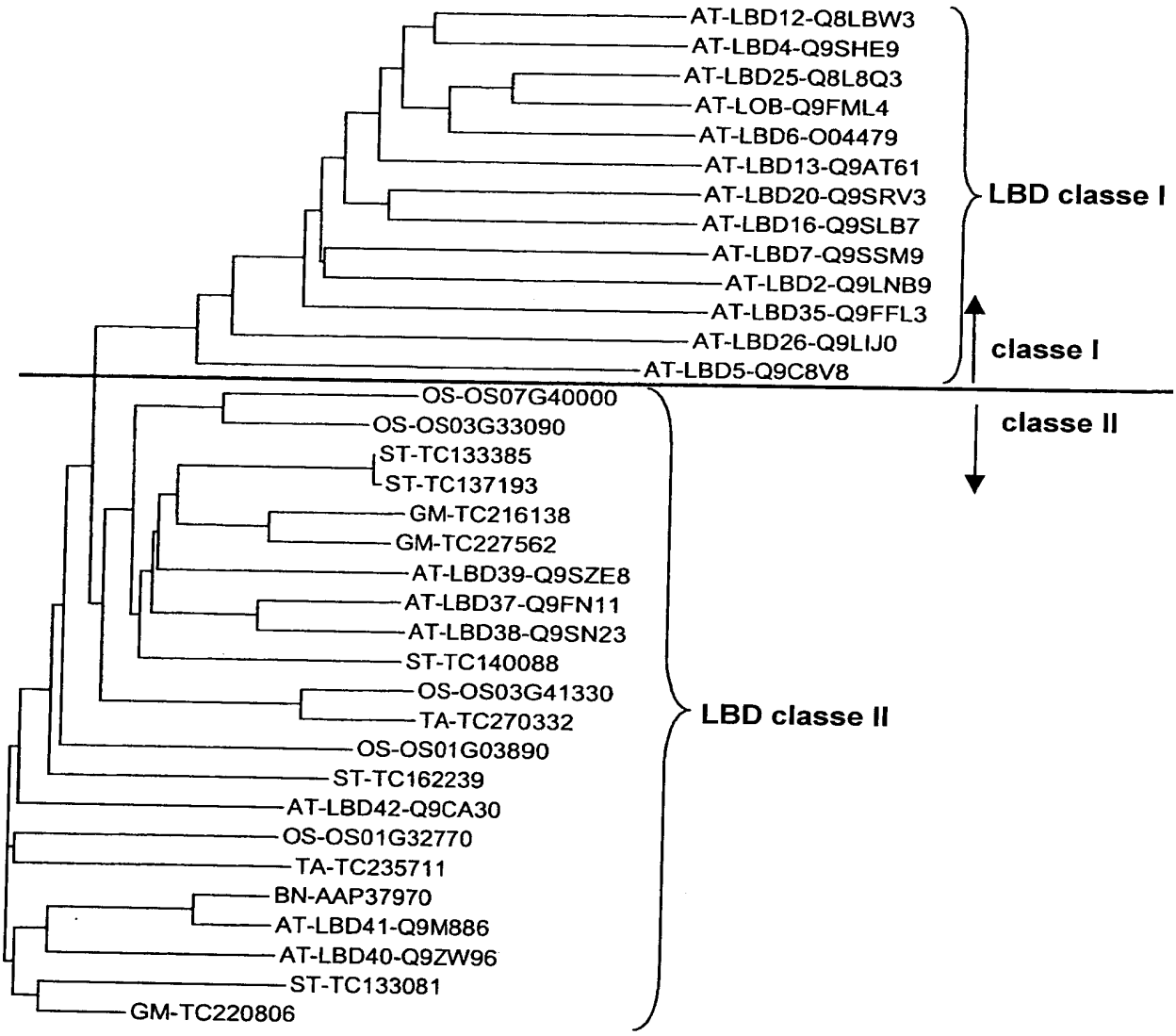

AT-LBD12-Q8LBW3
AT-LBD4-Q9SHE9
AT-LBD25-Q8L8Q3
AT-LOB-Q9FML4
AT-LBD6-O04479
AT-LBD13-Q9AT61
AT-LBD20-Q9SRV3
AT-LBD16-Q9SLB7
AT-LBD7-Q9SSM9
AT-LBD2-Q9LNB9
AT-LBD35-Q9FFL3
AT-LBD26-Q9LIJ0
AT-LBD5-Q9C8V8
LBD classe I
classe I
classe II
OS-OS07G40000
OS-OS03G33090
ST-TC133385
ST-TC137193
GM-TC216138
GM-TC227562
AT-LBD39-Q9SZE8
AT-LBD37-Q9FN11
AT-LBD38-Q9SN23
ST-TC140088
OS-OS03G41330
TA-TC270332
OS-OS01G03890
ST-TC162239
AT-LBD42-Q9CA30
OS-OS01G32770
TA-TC235711
BN-AAP37970
AT-LBD41-Q9M886
AT-LBD40-Q9ZW96
ST-TC133081
GM-TC220806
LBD classe II

FIGURA 3

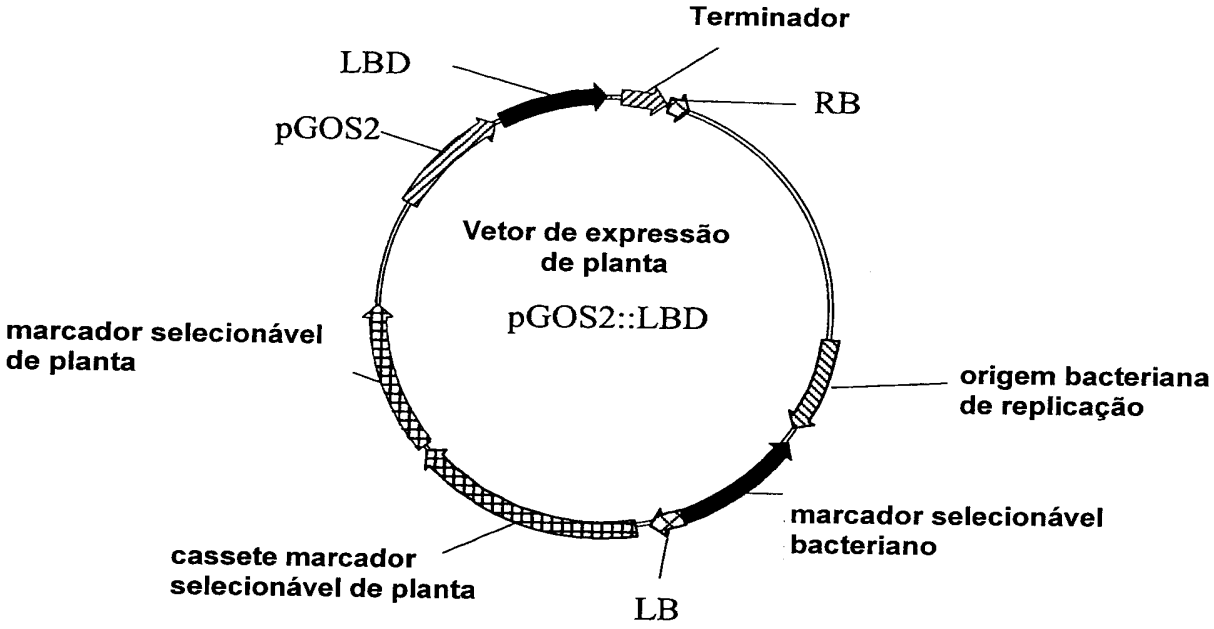

Terminador
LBD
RB
pGOS2
Vetor de expressão de planta
pGOS2::LBD
marcador selecionável de planta
origem bacteriana de replicação
marcador selecionável bacteriano
cassete marcador selecionável de planta
LB

FIGURA 4

SEQ ID NO: 1, Sequência codificadora de proteína LBD de Arabidopsis thaliana
ATGAGCTGCAATGGTTGCCGTGTTCTCCGGAAAGGTTGCAGCGAGAATTGTATCCTCCGGCCATGT
ATTCAATGGATTGAAACCGCCGATGCTCAAGGCCACGCCACCGTCTTCGTCGCTAAATTCTTCGGC
CGTGCTGGTCTCATGTCCTTTATCTCCGCTGTTCCGGATTCTCAACGTCCTGCTTTGTTTCAGTCG
TTGCTCTACGAAGCTTGTGGAAGAACTGTCAATCCAGTTAACGGAGCAATCGGAATGTTATGGACT
GGAAACTGGAATATCTGTCAAGCGGCTGTTGAAACAGTGCTTCGCGGCGGTTCTTTAAGACCGATC
CCGGAGCTTCTCACTCACGGCGGCGGTTTCGCTGGCTTTCCTTCGCCTACATCTGAAGAAGCATCT
GAGATCTGCACCGAAATGTTGAATCTCCAGCAAAATGATTCCACCGATCGTAACATCTATCATCAT
TCACGATTCTCAAGCTCTAGATCTAGATCTACTATGGATTCTTCTTCTCCGACGAAACGTAAGAGA
TTATCATCGGAAGACCAACCATCTTCGGAGCTTGATCTATCTCTCATCCCTAATTTTCCCATTAAG
CAAGCAACACCTTCTTCTACACGGCGGCGATCAGTAACACCGTCGATGAACTCAGAGGACTCCGGG
ACGACGACGACTACGACGGCGTTTTGTGACAAGGGTGATGTGTACGGTAACGGAGGAGGAGAAACG
ACCAAGTTGCTTAACCTTTTTGTTTAA

SEQ ID NO: 2, sequência de proteína LBD de Arabidopsis thaliana , Q9FN11
MSCNGCRVLRKGCSENCILRPCIQWIETADAQGHATVFVAKFFGRAGLMSFISAVPDSQRPALFQS
LLYEACGRTVNPVNGAIGMLWTGNWNICQAAVETVLRGGSLRPIPELLTHGGGFAGFPSPTSEEAS
EICTEMLNLQQNDSTDRNIYHHSRFSSSRSRSTMDSSSPTKRKRLSSEDQPSSELDLSLIPNFPIK
QATPSSTRRRSVTPSMNSEDSGTTTTTTAFCDKGDVYGNGGGETTKLLNLFV

SEQ ID NO: 3, prm009067 (iniciador avançado)
GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGAGCTGCAATGGTTGC

SEQ ID NO: 4, prm009068 (Iniciador reverso)
GGGGACCACTTTGTACAAGAAAGCTGGGTACTAACTCTGAGAAAACCGCC

SEQ ID NO: 5, motivo conservado 1
MSCNGCRXLRKGCX

SEQ ID NO: 6, motivo conservado 2
QXXATXFXAKFXGR

SEQ ID NO: 7, motivo conservado 3
FXSLLXEAXG

SEQ ID NO: 8, DP(V/I)YG assinatura
DP(V/I)YG

SEQ ID NO: 9, motivo Cys conservado
CXXCXXXXXXCXXXC

SEQ ID NO: 10, sequência de promotor GOS2
AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAATAT
AAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCACCTACT
TTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGTAATTAAGT
GGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGTTAAATTATTCG
AGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTTCTAGCTGAACTCAATGGGT
AAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTCTGAACGTATTGGCAAAGATTTAAA

FIGURA 5

CATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACATCATTAAGGACATGTCTTA
CTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATTTTTATTATTTATCTTTTTTCG
ATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAAT
ACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTAGGTAGCAATATC
TGAATTCAAGCACTCCACCATCACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTT
ACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAAAAAAGAATT
TTGCTCGTGCGCGAGCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCCACA
GAACAACCCACAAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAG
GCTTTGCGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGCATCCTCCTTCTCCCATCTATAA
ATTCCTCCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAAG
GACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTCGATCCATATCTTCCGGTCGAGTTCTTGGT
CGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCTCCCTTCGGTTGTTCTTGGATTT
ATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCC
TGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTCGGTTTGATTAGTAGT
ATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTAGGGATCGGAATCTTGCGATT
TTGTGAGTACCTTTTGTTTGAGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAGTAC
GGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTC
CCTATTGAACAAAAATAATCCAACTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTT
AAGCCTGTCCAAAATTTCGCAGCTGGCTTGTTTAGATACAGTAGTCCCCATCACGAAATTCATGGA
AACAGTTATAATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTT
TTTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATG
CTTTTATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGA
AGAACTTATCCGATTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATT
CATTTGGATTATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAA
CTGTCCTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGT
AGAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAATCG
GGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCTTGCCACT
TTCACCAGCAAAGTTC

SEQ ID NO: 11: Os-Os01g03890, Oryza sativa

MRMSCNGCRVLRKGCSDNCAIRPCLQWIRSPDAQANATVFLAKFYGRAGLINLITAGPEHVRPAIF
RSLLYEACGRMLNPVYGSVGLLWSGNWQLCQSAVESVLRGMPIAQPPPSATAVPPLPTCDIRHVGA
RRGDVHGAAAGPVADLHRLDISSRAKFKRPGGGAAAAHRSDHAAFELVFSKPAAAMAVDVIRQAQP
LNWAPGALSHESASHDAAPPESEGHSNDTADTVDGSHVSQSEPEPRATSAATEVHDAGLDLTLGLP
PPPPPVQKTEPADSDGGSQQQHDHRKEKPVELGLAISTKVAAQ

SEQ ID NO: 12, Os-Os01g32770, Oryza sativa

MRMSCNGCRVLRKGCSEGCTIRPCLQWIKTPEAQANATVFLAKFYGRAGLLNLLAAGPDHLRPAVF
RSLLYEACGRIVNPIYGSVGLLWSGQWQACQAAVEAVLKGDPVVQVSSEAAAAAQATPPLRAYDIR
HVSKDAEADAAANLLRVARGGRTRFKRASSSSNSKHGAKLAGAAAAKRAASPSSSSPTHETEPEAV
VVVGDHDDDHHHPALSHEVHEESAGSHDHDDDDHVDDGDNNDMAIADVTPPRAGSEDTEVETGSHV
SQAEQSPVPVEHEEGEEEEVGLELTLGFQPLVVRASRRPSSAEARCDLSGLSAESSRIGLRLELPA

SEQ ID NO: 13, Os-Os03g33090, Oryza sativa

MSCNGCRVLRKGCSDGCVLRPCLQWIDAADAQGHATVFVAKFFGRAGLLSFISAVPEAQRPALFQS
LLYEAAGRTINPVHGAVGLLWTGNWPLCQAAVETVLRGGAIGPLPELGGACGGAGGDLYGAAKRNG
GWSTFSTAKRVRKAEVPEAPSCDLGLCLSPGSPPAVGERKPALRPGTPSMSSDESGTTTGGERDPV
LLNLFV

FIGURA 5 CONT.

SEQ ID NO: 14, Os-Os03g41330, Oryza sativa
MSCNGCRVLRKGCSDACVLRPSIEWIDGAQPQANATVFVAKFFGRAGLVASLAAVPLHHRPALFRS
LLYEACGRTINPVSGAIGLMWTGNWDLCQAAADAVLRGDSLRALSAVPAAFTDRDMAGLYGNVGAA
AGTSSSSPDNDNSSASAAPRRKRPRNNGAGAGVGQQQLPHAVAAVLQSCELDLCLTPVSPPAVQLV
GGGGGASDEHSTTTCEEASDGDGAGAPTLLNLFS

SEQ ID NO: 15, Os-Os07g40000, Oryza sativa
MSCNGCRVLRKGCSEGCVLRPCLQWIDGAEAQGHATVFVAKFFGRAGLMSFLTAVPEPQRAAVFQS
LLYEAAGRTINPVGGAVGLLSGGSWHLCQAAVDTVLRGGGIQPLPDQVDAAAAGGRDVFASTARRA
MGGCSTFSTAKRSTTTTTTKNPGTPHDAAAAAPQPEPSCDLGLWLSPGSPPAPGDRRSGGRRADTP
SMNSEGSVTTCGVVGDGEREPELLNLFV

SEQ ID NO: 16, Nb-TC9404, Nicotiana benthamiana
MSCNGCRVLRKGCSENCILRPCLQWIETAESQGHATVFVAKFFGRAGLTSFISAVPENQRPSLFQS
LLYEAAGRTVNPVNGAVGLLWTGNWHVCQAAVETVLRGGALRPIPDFLGASLDIDESSECTDIFKL
QHPSQNNIQPTMQKRRRFPEEASNVLQLADLDLSLTPGFQSQKVFNHALPENRRPATPSMNSEESG
TTTCFESGTVIGNHQGNEPKLLSLFN

SEQ ID NO: 17, Gm-TC227562, Glycine max
MSCNGCRVLRKGCSESCILRPCLQWIETAEAQGHATVFVAKFFGRAGLMSFISNVPENQRPALFQS
LLFEACGRTVNPVNGAVGLLWTGNWHVCQAAVETVLRGGTLRPMPELMGLDAAAAAAAAADETSEA
EVGCTDTWRIRNPNTNCRFTSSRSSSGGGGGKRKRSEELAKLQEQQPNLDLRLTPVFLQKEESRRP
ESPSMTSEESGTTTENRFGDQWSHWKVLNLFI

SEQ ID NO: 18, Gm-TC216138, Glycine max
MSCNGCRVLRKGCSESCILRPCLQWIDTPEAQGHATVFVAKFFGRADLMSFISNVPEPQRPALFQS
LLFEACGRTVNPVNGAVGLLWTGNWHVCQAAVETVLRGGTVRPMPELFGLDAPTPTTDDASEGEVT
CTDKWRVRDPNPNFRFPGSRSDKGSSGGKRKRFEELAKLQTATTDLNLRLTPSFLQNAANFGCRQE
IRWPGSPSMNSEESGTTTACLESGIGDHYAPDGDRKVLNLFI

SEQ ID NO: 19, Hv-TC147776, Hordeum vulgare
MSCNGCRVLRKGCNDACMLRPCLLWIEGADAQGHATIFAAKFFGRAGLMSFLTAVPESQRPAVFQS
LLYEAAGRTINPVGGAVGLLWAGSWHLCEAAVQTVLRGGAIRPLPELAGGVPEGGVGGSDLFASSS
RRAVVGCSTYSMAKRVTPRKTWAPEAASHHQEPSCDLGLFLTPGSAAAAEGERRARRAGTPSMSSD
GSVTTTAGAGADGDKEPELLNLFV

SEQ ID NO: 20, Sb-TC104758, Sorghum bicolor
MSCNGCRVLRKGCSEACVLRPCLQWIDAADAQGHATVFVAKFFGRAGLLSFISAVPDAQRPALFQS
LLYEAAGRTINPVHGAVGLLGTGNWHLCQAAVDTVLRGGAIGPLPELGITGSAGDLYGPPGAGTGH
GGGGGKRAGGWSTFSTAKRVRKAAATGAPPPPPAAEPEASCDLRLCLLSPGSPPPPAGDRRAPFMR
PGTPSITSDDDSVTTTTTTTTTGGGGDREPVLLNLFA

SEQ ID NO: 21, Aq-TC18561, Aquilegia sp.
MSCNGCRVLRKGCSESCILRPCLQWIESPEAQGHATVFIAKFFGRAGLMSFISSVPESQRPALFQS
LLFEACGRTVNPVNGAVGLLWTGNWHICQAAVETVLRGGTLRPINEVLGASLMMMPESADEMSEVI
TTTQDPTYSRPKVMHIKKKEIIHENVNKIMIHDLDLSLTSGSSSPRKTTKRSTTPSTEESVTTSLD
SGAGVWRDDHYLRRGSGEDNNNNGDQKLLNLFV

FIGURA 5 CONT.

SEQ ID NO: 22, Vv-TC60668, Vitis vinifera
MSCNGCRVLRKGCSESCILRPCLQWIESAESQGHATVFVAKFFGRAGLMSFISAVPENQRPALFRS
LLYEACGRTVNPVNGAVGLLGTGNWHVCQAAMETVLRGGTLRAMPELLNGVSTPECDEASEAEVTC
TDMFKLRDPNSSSRSKVPKRKRLEEASKVQCTDLDLRLTPGLTGKISSRGLGAMSYLPETRRPGTP
SMNSEESVTTTCFETPHSPEGERKLLNLFL

SEQ ID NO: 23, Lj-TC15459, Lotus japonicus
MSCNGCRVLRKGCSESCILRPCLQWIDTAEAQGHATVFVAKFFGRADLMSFISNVPQPQRPSLFQS
LLFEACGRTVNPVNGAVGLLWTGNWHVCQAAVETVLRGGTLKPLPELLGLDAPSHVPTAADDASEG
EVTYDTWRIRDPNPNARLTNSRSSAGKRQRPEELAKLQTTTDLNLRLTPGFLQNVSSYGCKREARR
PGSPSMNSEESVTTTACLGSGIGDHCSRDEDRKVLKLFV

SEQ ID NO: 24, Gh-TC30552, Gossypium hirsutum
MSCNGCRVLRKGCSESCILRPCLQWIESPEAQGHATVFVAKFFGRAGLMSFISAVPESQRPALFQS
LLFEACGRTVNPVNGAVGLLWNGNWHVCQAAVETVLRGGSLRSMPELMAPTPASDEASEATCTDMW
KLRETSTNLNSNCRFSNSRSRVSPKRKRVEEEFKKLQPSDLDLGLTPSFTGKRVTDNRRPGTPSMN
SEESVTTTCFESVFADQQGQGSRGTDKKLLKLFV

SEQ ID NO: 25, Ta-TC235711, Triticum aestivum
MRLSCNGCRVLRKGCSEDCSIRPCLQWIKSPEAQANATVFLAKFYGRAGLMNLINAGTDDSLRPGI
FRSLLYEACGRIVNPIYGSVGLLWSNNWQMCQAAVEAVLSGKPIVQVSSEDAAADRTPPLKAYDIR
HVSTSPAADGRLHKVAKPGRTRFKRASSASSHHNPSSDSNNKPKPKPQPQTRAPTAEEERDRQHRK
EMEEGAFQRAPSHESSDSRHEDPVEPHCQQEASADTEAEAGSHVSQAEQEQSTEPAAEHAEEVEKE
DEELGLELTLGFAPVAARPAEFT

SEQ ID NO: 26, St-TC133081, Solanum tuberosum
MRMSCNGCRVLRKGCSKSCSIRPCLQWIKTPDSQSNATVFLAKFYGRAGLMNLINAGPDHLRPAIF
RSLLYEACGRIVNPIYGSVGLLWSGNWQLCQNAVEAVLKGTPITPIASEIAVNNNGPPLKLPYDIR
HINKDENSTKSSELHRVRTRCRFKRSGANTKATKSNPACSGSGDEFAHEKVNGSTSHESSLSHQSE
EEAAAAAAVALNVECESPEMAEVEDSAKADGEVELELTLGFSSLGTVEGKPKETKRNKGVQLVDGA
GECKIELGLH

SEQ ID NO: 27, Mt-TC107091, Medicago truncatula
MRMSCNGCRVLRKGCSENCSIRPCLQWIKSPESQANATVFLAKFYGRAGLMNLVNAGPEHLRPAIF
RSLLYEACGRIVNPIYGSVGLLWSGSWQLCQAAVEAVLKGAPITPITSEAAANGRGPPLKAYDIRH
VSKEENSAASTELTQQRVKPRSRKRSSLAKLKIQEEEKSSNDNKGVESGSTEPTGSVEPVEEVMNR
SVSHESSISHQSEAVAVVDGESKDSESETSMILFRDEPELDRKMKPADRTGENGEEKVGLELTLGL
EPVSRVYHVVPVKKRRVVLKDCGVGSWNVELGLQYPA

SEQ ID NO: 28, Hv-TC147808, Hordeum vulgare
MRLSCNGCRVLRKGCSEDCSIRPCLQWIKSPEAQANATVFLAKFYGRAGLMNLINAGTDDSLRPGI
FRSLLYEACGRIVNPIYGSVGLLWSNNWQMCQAAVEAVLSGKPIVQVSSEDAAADRTPPLKAYDIR
HVSTSPAADGRLHKVAKPGRTRFKRASSASSHHNPSSDSNNKPKPQPRPPTAEEELDRQHRKEMEE
GAFQRAPSHESSDSRHEDPVEPHSQQEASADTEAEAGSHVSQAEQEQEQSTEPAADHAEEVEKDDE
ELGLELTLGFAPVAARPAGCHLSVRTAAEPAFVGLRFL

FIGURA 5 CONT.

SEQ ID NO: 29, Vv-TC55931, Vitis vinifera
MRMSCNGCRVLRKGCSDDCSLRPCLEWIKTPESQANATVFLAKFYGRAGLMNLLNAGPEHLRPAIF
RSLLYEACGRIVNPIYGSVGLLWSGNWNHCQNAVDAVLGGSQIMEVPCDTAAIRPILPLKAYDIRH
LSKDANSPHDLHKAKTRSRFKRCGARSKPHLDSAEFCNPGWNQFSPVDYNRAPSHDSSLIQQETDS
MFSVETVEASLATPPKPNQLLKSDGQTDKTEAGLELTLGFYPISRPHACKLELGP

SEQ ID NO: 30, St-TC162239, Solanum tuberosum
MRMSCNGCRVLRKGCSENCSLRPSLQWIKSPDSQANATVFLAKFYGRAGLINLINAGPDHLRPAIF
RSLLYEACGRIINPVNGSVGLMCSGNWQRCQEAVESVLKGSTIMQVPINNDDDNNNNNNNNTDQIV
PLKGCDIRHVSKNNSTNSDNQQVIKTRKGRLKRKGTFDSVASASAEAEAAEEYLLKNQPPLKFSIA
GWDQFEEIDDELIKRAASHDSFSVETVELTLLGEPG

SEQ ID NO: 31, Pta-TC69225, Pinus taeda
MRVSCNGCRVLRKGCSDSCELRPCLQWIKTSESQAQATVFLAKFYGRSGLMNLIKSGPEFLRPTLF
RSLLYEACGRIVNPIFGSVGLIWSGNWHICQAAVESVLKGYPLQPLHPLNSLEIASPPPAAAAPQS
MKAVADKLRNVRSQGRFKRPLYRKPSRDVKNLNLHALKDEEAAGILTGSQAYNLRWRRCYPKEFQL
INAQPVNNFLSIDSSSDSDDLIRLKTGFESVKSDCNPIENRETRALSEGISEGPALSEGIENQNEC
PLQNLFQQNDNNWVENNIMRTELTDERNLGLELTLGSSMGETTQSHRKQLGSNCMDRPPLYRSLDL
SLSQ

SEQ ID NO: 32, Pta-TC67269, Pinus taeda
MRMSCNGCRVLRKGCSETCIIRPCLQWIKNEESQANATVFLAKFYGRAGLMNLISAGPEHLRPAIF
KSLLYEACGRMVNPIFGSVGLLCSGNWQLCQAAVESILKGCPIHPILVDSSGTQMPTERSNTNPSV
RSTLMETITNELHKVKTKGRFKRRKSTVKDVSLNNSDAEDLGASQVFNLRWRRCYPGESQTPALEG
IPNHLVSSDSSCESLEHFDDLRNHKEADMEGHMSKAMDMGPVNIKCCSDQNDHKATMESVPEVSYT
QAPDLEDMQAIKHEYESISSDVTDVKLELTLSSQGSATHAAEKHHTNLHWNSCMEHSDEPTITYLG
LGLTQA

SEQ ID NO: 33, Gm-TC220806, máxima glicina
MRMSCNGCRVLRKGCSEDCSIRPCLQWIKNPESQANATVFLAKFYGRAGLMNLINAGPENLRPAIF
RSLLYEACGRIVNPIYGSVGLLWSGSWQLCQAAVEAVLKGEPITPITSEAAANGRAPPLKAYDIRH
VSKDQNSANETPKTKTRSRFKRTSGTLIKPKASKGTGFVPVEPEMANPDCES

SEQ ID NO: 34, Ta-TC270332, Triticum aestivum
MSCNGCRVLRKGCSDACVLRPSIEWIDGAQPQANATVFVAKFFGRAGLVASLAAVPLHHRPALFRS
LLYEACGLTINPVSGAIGLMWTSNWDLCQAAAEAVLRGDSLRSLSAVPAAFTERDMAGLYGNVGTN
TGSSSSLHSSPENSTSAPARKRSKNNCGAAVGQQVKLPGPGPVLQSCELDLCLTPLSSPLAGGRRG
GASDEYSTTTCCEEASGDAAEAGAPPLLNLFN

SEQ ID NO: 35, Aq-TC18329, Aquilegia sp.
MRMSCNGCRVLRKGCSENCSIRPCLQWIKTPESQANATVFLAKFYGRAGLMNLINSGPEHLRPAIF
RSLLYEACGRIVNPIYGSVGLLWSGSWQLCQAAVEAVLKGTPIMQISSEVAAISPSPPLKAYDIRH
VSKDENSSAASQELHKVKTRSRFKRSGAGKAANNKQVNESNKVHTYLDRAPSHESSVSHQTELNNG
EGCSRETESMFSVETVEASLVNRADPEPVPKFDNFNGSDADDNNNDDVELELTLGFEPSRREKSMS
PKKWRDEINNSYTDTCKIEVGIDYAV

FIGURA 5 CONT.

SEQ ID NO: 36, St-TC137193, Solanum tuberosum
MSCNGCRVLRKGCSDNCILRPCLQWIETPEAQGHATVFVAKFFGRAGLMSFISAVPENQRPALFQS
LLYEAAGRTVNPVNGAVGLLWTGNWHVCQAAVETVLRGGALRPISEFLGASVEIDEVSDCTDVFKL
QDPSLNMRPKMQKRRRFPEETSMLADLDLSLTPGFNQKVYNSHPLAENHRRPGTPSMNSEESGTTT
CFESTAVIGEEPKLLSLFN

SEQ ID NO: 37, St-TC133385, Solanum tuberosum
MSCNGCRVLRKGCSDNCILRPCLQWIETPEAQGHATVFVAKFFGRAGLMSFISAVPENQRPALFQS
LLYEAAGRTVNPVNGAVGLLWTGNWHVCQAAVETVLRGGALRPISEFLGASVEIDEVSDCTDVFKL
QDPRKMQKRRRFPEETSMLADLDLSLTPGFNQKVYNSHPLPENHRRPGTPSMNSEESGTTTCFEST
AVIGEEPKLLSLFN

SEQ ID NO: 38, St-TC140088, Solanum tuberosum
MSCNGCRILRKGCNENCILRQSLQGIESPHAQANATVFVAKFFGRAGLMSFLTSVSESQRPALFQS
LLFEACGRTVNPVHGAVGLLWTGNWHVCQSAVETVLKGGVLRPLPEFSGVTASPEFDSEANDVDLF
RSQSSNFQSKRKIIDEVSEDLDLGLTSGLSTMVTGGKLNRRTEKRRAATPSLNSDESDTTTLESGV
VYHQNPQQGNETKLLRLFF

SEQ ID NO: 39, Pa-TC14656, Picea alba
MSATDKNITERKEQQRKMSCNGCRVLRKGCSDNCILRQSLQWIESPQSQAHATVFVAKFFGRAGLM
TFIATVSEPERPDLFKSLLYEACGRTVNPVHGAVGLMWTGKWQICQAAVQTVLSGGSLDRTPASPW
FPLTDANGDRNRMSAGDQWPNETDVGNYPGDDDVCKIYELDACVDSRRVRQKIDGSNNNNNEIAST
DADKLGQRGSGIDGQKNDIKLDLSLNCSSKSKSSNLKQQRVSSPSTNSVHSEGSVNRLKISNPLPS
STMASQQKFLDLL

SEQ ID NO: 40, Pta-TC59178, Pinus taeda
MTATTGAGNSGRDLSVFIQRKQEADGRKRMSCNGCRVLRKGCSDTCILRPCLQWIESPEAQGHATV
FVAKFFGRAGLMSFISAVPDHQRPALFQSLLYEACGRTVNPVNGAVGLLWNGNWQLCQAAVETVLK
GGTLRPPQSSSVIPTCSTRPSLMNSFSANFRVNSNVNNINSSGSRLSLEVGKARPVNEMREMYTVK
PEVDFPLNRPQESCNPHQPWAAKVKSEPHDLTLCEAFEAVRHSQTIIHSPPSEIAAAPRRVRQRIE
GINEVNYRDLSNSNSDDVKCLGVFDIHPNNTMEMKANLDLTLNFEASSSSHIKRRVSSPSIVSVNS
EGSVTSLESGSLRVPNQKDSSAAKEHKVLDLFP

SEQ ID NO: 41, Pta-TC67974, Pinus taeda
MSATENHITERKEQERKMSCNGCRVLRKGCSENCILRQSLQWIESPQSQAHATVFVAKFFGRAGLM
TFIAAVSEPERPALFKSLLYEACGRTVNPVHGAVGLLWTGKWQICQAAVQTVLSGGSLDRTPAGTC
FPSTDANGDRNGMSDGDQWPNEMDGNFPGDDDVCKIYELGACDDSRRVRQKIDGNNNNINELMSTD
GEKLGQRGSGIDGQKNDLKLDLSLNCPSKSKSSNLKQQRVSSPSTNSVHSEGSVNRLKSSSPPPSS
TMAAEQKLLDLL

SEQ ID NO: 42, Le-TC178827, Lycopersicon esculentum
MSCNGCRVLRKGCSDNCILRPCLQWIETPEAQGHATVFVAKFFGRAGLMSFISAVPENQRPALFQS
LLYEAAGRTVNPVNGAVGLLWTGNWHVCQAAVETVLRGGALRPISEFLGASVEIDEVSDCTDVFKL
QDPSLNMRPKMQKRRRSPEETSMLDLSLTPGFNQKVYNSHPLPENHRRPGTPSMNSEESGTTTCFE
SSAVIGDHQGKEPKLLSLFN

FIGURA 5 CONT.

SEQ ID NO: 43, Pt-III.589, Populus tremuloides
MSCNGCRVLRKGCSENCILRPCLQWIESPEAQGHATVFVAKFFGRAGLMSFISAVPENQRPSLFQS
LLFEACGRTVNPVNGAAGLLWTGNWHVCQAAVETVLRGGTLQPMPELLTGGGGSPSPSSDEASEVE
VACTDIWKLQDPNPNHHPRFSISRSRLSPKRKRTEEPVAVKVQHNDLDLGLTPSFSLKGFPYKQQI
RRPGTPSMNSAESVTTNTACFDSAGLGDEGGETKLLNLFL

SEQ ID NO: 44, Pt-V.543, Populus tremuloides
MSCNGCRVLRKGCSENCILRPCLQWIESPEAQGHATVFVAKFFGRAGLMSFISSVPEDQRPSLFQS
LLFEACGRTVNPVNGAVGLLWTGNWHVCQAAVETVLRGGTLRPMPDLLTGGGSPSSSPPSDEASEF
EVACTDIWKLQDPNPNPIHHSRFSNSRSRVSSKRKGTEEPMVVNMRHDDLDLLLTPSTSQKGFAEI
CRPGTPSMNSEESVTTNATCFDSAGFGDQYGNGGGETKLLNLFL

SEQ ID NO: 45, Pt-XIV.94, Populus tremuloides
MSCNGCRVLRKGCSETCVLRSCLHWIPTPEAQGNATLFLAKFFGRSDLISLISAVPESQRPVLFQS
LLFEACGRTVNPVNGAVGLLWSGNWHVCQAAVESVLAGETLRPLPGILTGVLAPNCDESSDSFSAA
AYAVHSMTWNQSKPFDKENNNQVVSDHHVNLCLARGRGGRDKRGRDAVSFYTGGESETTSFESSGA
DKNKLLNLFV

SEQ ID NO: 46, Pt-II105, Populus tremuloides
MSCNGCRVLRKGCSETCVLRSCLHWITNPEAQGNATLFLAKFFGRSDLMSLISAVPEAQRPALFQS
LLFEACGRTVNPVNGVVGLLWSGNWHMCQEAVETVLSGGALQPLPGILTGVLAPNCDESSDRFSAA
AYAVHNMARNQSRSFAKENNKEVVSDHRINLCLARGKGGRDKRGRDAVSFYTGGESETISFESSGG
DKKRLLNLFV

SEQ ID NO: 47, Pt-123.86, Populus tremuloides
MRMSCNGCRVLRKGCSENCSIRPCLQWIKSSESQANATVFLAKFYGRAGLMNLINAGPEHLRPAIF
RSLLYEACGRIVNPIYGSVGLMWSGSWQLCQAAVEAVLKGAPITPINSEAAVNGHGPPLKVYDIRH
VSKEENSAASNDANRARTRCRVRRVVKPKASKRACFGNGLASIAVDTLSARDELTRCTSHESSVSH
QSELAVLDGDSKESDESMVSVETAEASLLFPPNPESKNDCKAHDVASNGISGLELGLDLTLGLEPV
SRAHHVVPVKKRRVEAYGSGDVDTCKMELRLE

SEQ ID NO: 48, Pt-X180, Populus tremuloides
MSCNGCRVLRKGCNDNCTIRPCLQWIKSTDSQANATLFLAKFYGRAGLINLIEAAPQRLRPAIFSS
LLYEACGRIVNPVYGSVGMLWSGNWAQCQAAVDAVLKGLPIISIPSSDVPTPHLISSLNTYDIRHV
SRDQNSPELNKVKSRTRCKRSIGTRPSSLAEPTGRLNQGELGFEPVRESWLGHLGNGDSRIKDEDS
ILAAENVEDSLWSRVEPDQVFKFNGQSDDSDLSLELTLALVSE

SEQ ID NO: 49, Pt-XII.481, Populus tremuloides
MPFSALSFSSLPHRIILLLIGFLNSSPSLSRHFSISFRTEYTMRMSCNGCRILRKGCGDNCSIKPC
LQWIETPDSQANATLFLAKFYGRAGLMNLINACPQHLRPDTFKSLLYEACGRIVNPVSGSVGLLST
GSWQQCQAAVEAVLKGEPITQIASTDQLTLGGSDTRHVSREEDWSASDQPRKIKSKRQFNRSTSKR
KPSRTEAEHTCFEFMIGFDGLRSVSPDSVLSWRPSLGSDNEETDSIGSVETVEASKLDEPAEGSDL
DLDLTLGHY

SEQ ID NO: 50, Ck-LOB-DQ787782, Caragana korshinskii
MSCNGCRVLRKGCSESCILRPCLQWLENPEAQGHATVFVAKFFGRAGLMSFLSNVPETQRPALFQS
LLFEACGRTVNPVNGAVGLLWTGNWHVCQAAVETVLQGGTLRPLPEHTLLDAPPAATDEASEAEDA
CTGIWRIRDPNPNCRPNGSRSNNKVSTGGKRKRSEELVKIPAATNLDLRLTPIFLQKESESRRPGS
PSMNTEESVTTTATGLGDSWGHGGERKVLNLFI

FIGURA 5 CONT.

SEQ ID NO: 51, Bn-AAP37970, Brassica napus
MRMSCNGCRVLRKGCSDDCSIRPCLGWIKSPEAQANATVFLAKFYGRAGLMNLINAGPDHLRPGIF
RSLLHEACGRIVNPIYGSVGLLWSGNWQLCQSAVEAVMKGEPITEMATDAATSGQGPPLKMYDIRH
IAKDENSPAAAAAAGSTDRKRGKTRRAKRVAAVAKPAESGGGEASHHSSLSHQSEVVVAPHEGESK
ESESNNSEVMTFSPPAVQGSGEIKLDLTLGLEPVSRADHVVPVKKRKMGVFSTWQEESSCKTDPVL

SEQ ID NO: 52, Mt-ABE82505, Medicago truncatula
MSCNGCRVLRKGCSESCILRPCLQWIETPEAQGHATVFVAKFFGRAGLMSFISNVPEPQRPALFQS
LLFEACGRTVNPVNGAVGLLWTGNWHVCQAAVETVLRGGTLRPLPELTLLDAPLTATDEASEAEDG
GADAMWKIREPNFGLASSRSSNKVCSGSKRRRSEEFVKVPAAINLDLRLTPIFQQKAVEERRHGSP
SMTSEESVTTTACLETGIGDRWSHGGDRKVLNLFI

SEQ ID NO: 53, Mt-ABE78739, Medicago truncatula
MRMSCNGCRVLRKGCSEDCSIRPCLEWIKCPQSQANATLFLAKFYGRAGLINLINSGSENLRPAIF
RSLLYEACGRIVNPIYGSVGLLWSGSWHLCQAAVEAVLKGEPITPIDSEADENERGPPFKAYDIRH
VSKDENLEETKQVRTRSRFRRAMKPKPSKEVGSGSGSSSAKEVDRSKSQESSLSQPGKDSESGVSV
ETSILFHDESEFEPESQAKVKGRAEGRAEEGLSDGVGLELRLGIEPVSREEFMVPIKKRKIVLKDC
SASSKVELGLELSA

SEQ ID NO: 54, At-LBD38-Q9SN23, Arabidopsis thaliana
MSCNGCRVLRKGCSENCILRPCIQWIESPEAQGHATVFVAKFFGRAGLMSFISAVPESQCPALFQS
LLYEACGRTVNPVNGAVGLLWTGNWNVCQAAVETVLRGGSLKPIPELLNGGGFAGFPSPTSDEASE
ICTEMLNLRKADDSGDRNIYHHCRFSSSRSRSRSTASPPKRKRLSSEQQPSSELDLSLIPIYPIKT
LPFKEDTPSMYSEESVTTVSFQNNNAGDRYVRCGGGGGGATTKLLNLFA

SEQ ID NO: 55, At-LBD39-Q9SZE8, Arabidopsis thaliana
MSCNGCRVLRKGCSETCILRPCLQWIESAESQGHATVFVAKFFGRAGLMSFISSVPELQRPALFQS
LLFEACGRTVNPVNGAVGMLWTRNWHVCQAAVETVLRGGTLRPISDLLESPSLMISCDESSEIWHQ
DVSRNQTHHCRFSTSRSTTEMKDSLVNRKRLKSDSDLDLQVNHGLTLTAPAVPVPFLPPSSFCKVV
KGDRPGSPSEESVTTSCWENGMRGDNKQKRNKGEKKLLNLFV

SEQ ID NO: 56, At-LBD40-Q9ZW96, Arabidopsis thaliana
MRMSCNGCRVLRKGCSENCSIRPCLQWIKSAESQANATVFLAKFYGRAGLMNLLNTGPDHLRPAIF
RSLLYEACGRIVNPIYGSVGLLWSGNWHLCQAAVEAVMRGSPVTPIACDAAVTGQAPPFNNKLCDI
RHVSSRDENVKRRSRGACKEERNVRSLSHESSLSHESPVSSEETTTEEPKTWIGLELTLGLEPLAR
GNHVVVPMKKRKLERCGTSEDEDTCKIELGLVCSE

SEQ ID NO: 57, At-LBD41-Q9M886, Arabidopsis thaliana
MRMSCNGCRVLRKGCSEDCSIRPCLAWIKSPEAQANATVFLAKFYGRAGLMNLINAGPNHLRPGIF
RSLLHEACGRIVNPIYGSVGLLWSGNWQLCQDAVEAVMKGEPVKEIATDAATIGQGPPLKIYDIRH
ISKDDNSAAAATGSTDLKLAKTRRAKRVSTVAIQAESEGKSDEASHDSSLSHQSEIVAAHEGESKE
SESNVSEVLAFSPPAVKGSGEIKLDLTLRLEPVSRAYHVVPVKKRRIGVFGTCQKESTCKTELML

SEQ ID NO: 58, At-LBD42-Q9CA30, Arabidopsis thaliana

MRISCNGCRVLRKGCNQDCTIRPCLQWIKSADSQANATLFLAKFYGRAGLLNLIESGPDHLRPAIF
RSLLYEACGRIVNPVDGSVGLMWSGNWAQCQAAVDAVLNGLPITHTPLPSASASHQIIPPHRTYDI
RHVAKDPTTGGDSSENLATRVNANKSKTQTGRFKRESVNQLGECSHDMWQLPSSSATHGYGHFTLE
NVESRREAPFNQSSPNLGFDDQVDINEVGLELRLG

FIGURA 5 CONT.

SEQ ID NO: 59, Os-OS01G03890

ATGCGGATGAGCTGCAACGGGTGCCGCGTGCTGCGCAAGGGGTGCAGCGACAACTGCGCCATCCGC
CCCTGCCTGCAGTGGATTCGCAGCCCCGATGCGCAGGCCAACGCCACCGTCTTCCTCGCCAAGTTC
TACGGCCGCGCCGGCCTCATCAACCTCATCACCGCCGGCCCCGAGCACGTCCGTCCTGCCATATTC
CGGTCGCTGCTGTACGAGGCTTGCGGGCGGATGCTGAACCCGGTGTACGGATCCGTCGGCCTCCTC
TGGTCGGGGAACTGGCAGCTTTGCCAGTCCGCCGTCGAGTCCGTCCTCCGCGGCATGCCCATCGCG
CAGCCGCCGCCGTCCGCCACCGCCGTGCCCCCGCTCCCCACCTGCGACATCCGCCACGTCGGCGCC
AGGAGAGGCGACGTCCATGGCGCCGCGGCCGGCCCGGTTGCCGACCTCCACCGGCTTGACATCAGC
TCGCGCGCCAAGTTCAAGAGGCCCGGCGGCGGCGCCGCCGCCGCCCACAGGTCCGACCATGCCGCC
TTCGAGCTCGTCTTCTCCAAGCCCGCCGCCGCCATGGCCGTCGATGTGATAAGGCAGGCGCAGCCG
CTCAACTGGGCGCCCGGGGCGCTCAGCCACGAGTCGGCGAGCCACGACGCCGCGCCGCCGGAGAGC
GAGGGCCACAGCAACGACACGGCCGACACCGTGGACGGCTCCCACGTCAGCCAGTCGGAGCCGGAG
CCGAGAGCGACGTCGGCGGCGACCGAGGTGCACGACGCCGGCCTAGACCTCACATTAGGTCTACCG
CCGCCGCCGCCGCCGGTGCAAAAGACCGAGCCGGCGGACAGCGACGGCGGCAGCCAGCAGCAACAC
GATCATCGGAAGGAGAAGCCGGTGGAGCTCGGCTTGGCGATCTCAACTAAAGTAGCAGCTCAATGA

SEQ ID NO: 60, Os-OS01G32770

ATGCGGATGAGTTGCAACGGTTGCCGGGTGCTGCGCAAGGGATGCAGCGAAGGGTGCACCATCCGT
CCGTGCCTGCAGTGGATCAAGACCCCCGAGGCGCAGGCCAACGCCACCGTCTTCCTCGCCAAGTTC
TACGGCCGCGCCGGGCTTCTCAACCTGCTCGCTGCCGGGCCTGACCACCTCCGCCCCGCCGTCTTC
CGCTCGCTCCTCTACGAGGCCTGCGGCCGCATCGTCAACCCGATCTACGGCTCCGTCGGCCTGCTC
TGGTCGGGCCAGTGGCAGGCGTGCCAGGCCGCCGTCGAGGCCGTCCTCAAGGGCGACCCCGTCGTG
CAGGTCTCCTCTGAGGCCGCCGCCGCCGCCCAGGCCACCCCGCCGCTCAGGGCCTACGACATCCGC
CACGTCTCCAAGGACGCCGAGGCCGACGCCGCCGCCAACCTGCTCCGCGTCGCGCGCGGCGGCCGC
ACCCGCTTCAAGCGCGCCTCCTCCAGCTCCAACTCCAAGCACGGCGCCAAGCTTGCCGGAGCAGCC
GCCGCCAAGCGCGCCGCGTCCCCCAGCAGCAGCAGCCCGACGCACGAGACGGAGCCGGAGGCGGTG
GTCGTCGTCGGTGACCATGACGACGACCACCACCACCCTGCGCTGAGCCACGAGGTCCACGAGGAG
TCGGCTGGCAGCCATGACCACGACGACGACGACCATGTGGACGACGGTGACAACAACGACATGGCC
ATCGCCGACGTGACGCCGCCGCGGGCGGGATCGGAGGATACCGAGGTCGAGACCGGCTCACACGTG
AGCCAGGCCGAGCAGAGCCCCGTGCCGGTGGAGCACGAAGAGGGGGAGGAGGAGGAGGTCGGGCTG
GAGCTCACGCTCGGGTTCCAGCCGCTCGTCGTGCGCGCCTCGAGGAGGCCTTCGTCGGCGGAGGCG
CGCTGCGACCTTAGCGGCTTGAGCGCGGAGTCGAGTCGCATCGGCCTGCGGCTCGAGCTGCCAGCG
TGA

SEQ ID NO: 61, Os-OS03G33090

ATGAGCTGCAACGGTTGCCGGGTGCTGCGGAAGGGGTGCAGTGACGGCTGCGTGCTGCGGCCATGC
CTGCAGTGGATCGACGCCGCCGACGCGCAGGGCCACGCCACCGTCTTCGTCGCCAAGTTCTTCGGC
CGCGCCGGACTCCTCTCCTTCATCTCCGCTGTCCCCGAGGCGCAGCGACCTGCGTTGTTCCAGTCG
CTGCTGTACGAAGCAGCGGGGCGAACCATCAATCCGGTGCACGGCGCGGTGGGCCTCCTCTGGACG
GGGAACTGGCCCCTCTGCCAGGCCGCTGTCGAGACCGTGCTGCGTGGAGGCGCCATCGGCCCTCTG
CCGGAGCTCGGCGGGGCCTGCGGCGGCGCTGGCGGGGACCTCTACGGCGCCGCCAAGCGCAACGGC
GGCTGGTCTACCTTCTCCACTGCGAAGCGGGTGAGGAAGGCCGAGGTACCTGAAGCCCCGTCGTGC
GACCTAGGCCTGTGCCTCAGCCCCGGCTCTCCGCCGGCGGTGGGGGAGAGGAAACCAGCACTGCGG
CCGGGGACGCCGTCAATGAGCTCCGATGAGTCCGGCACCACAACCGGAGGCGAAAGGGATCCTGTG
CTGCTCAACCTTTTTGTCTGA

FIGURA 5 CONT.

SEQ ID NO: 62, Os-OS03G41330

ATGAGCTGCAACGGGTGCCGCGTGCTGCGGAAGGGGTGCAGCGACGCGTGCGTGCTGCGGCCGAGC
ATCGAGTGGATCGACGGCGCGCAGCCGCAGGCCAACGCCACCGTCTTCGTCGCCAAGTTCTTCGGC
CGCGCCGGCCTCGTCGCCTCCCTCGCCGCCGTCCCTCTCCACCATCGTCCAGCGCTGTTCCGGTCG
CTTCTGTACGAGGCGTGCGGGCGGACGATAAACCCGGTGAGCGGCGCCATCGGGCTCATGTGGACG
GGCAACTGGGACCTCTGCCAGGCCGCCGCCGACGCCGTCCTCCGCGGCGACTCCCTCCGCGCCCTC
TCGGCCGTGCCAGCCGCCTTCACCGACCGCGACATGGCCGGCCTCTACGGCAACGTCGGCGCCGCC
GCCGGGACCTCCTCCTCCTCGCCCGACAACGACAACTCCTCCGCCTCCGCCGCGCCCCGGAGGAAG
CGGCCCAGGAATAACGGCGCCGGCGCCGGCGTCGGTCAGCAGCAGCTGCCGCACGCGGTGGCTGCG
GTGCTCCAGTCCTGCGAGCTGGACCTCTGCCTCACGCCCGTCTCACCGCCGGCCGTGCAGCTGGTG
GGTGGGGGCGGCGGCGCGTCGGACGAGCACTCGACGACGACGTGCGAGGAAGCCAGCGACGGGGAT
GGCGCCGGGGCGCCCACGCTGCTGAACCTCTTCAGCTGA

SEQ ID NO: 63, Os-OS07G40000

ATGAGCTGCAACGGCTGCCGCGTGCTGCGGAAGGGGTGCAGCGAGGGGTGCGTGTTGCGGCCGTGC
CTGCAGTGGATCGACGGCGCCGAGGCGCAGGGCCACGCCACCGTGTTCGTCGCCAAGTTCTTTGGC
CGCGCCGGCCTCATGTCCTTCCTCACCGCCGTCCCCGAGCCGCAGCGCGCAGCGGTGTTCCAGTCG
CTGCTGTACGAGGCGGCGGGGAGGACGATCAACCCGGTGGGCGGCGCGGTCGGGCTGCTCTCCGGC
GGGAGCTGGCACCTCTGCCAGGCGGCGGTCGACACCGTGCTGCGCGGCGGAGGCATCCAACCGCTG
CCGGACCAGGTCGACGCCGCCGCCGCCGGTGGCCGCGACGTGTTCGCCTCCACGGCGAGGCGCGCC
ATGGGCGGGTGCTCCACGTTCTCGACGGCGAAGCGGTCGACGACGACGACGACGACCAAGAACCCC
GGTACTCCGCATGACGCCGCCGCCGCGGCGCCCCAGCCGGAACCGTCGTGCGACCTCGGCCTGTGG
CTCAGCCCCGGCTCCCCGCCGGCGCCCGGGGACCGGCGATCCGGCGGCAGGCGGGCCGACACGCCG
TCGATGAACTCCGAGGGATCCGTCACGACGTGCGGCGTCGTCGGCGACGGCGAGAGGGAGCCCGAG
CTGCTGAATCTTTTTGTCTAG

SEQ ID NO: 64, At-LBD38-BT002449

AAACATTGAAACACTTTGTCCCACTCTCTCTCTTTCTCTTTCTTGTACCAAAAGCTTTTTGAATCT
CCAAGATTATAGCAAAACCAAAGATAAAATACTAACTTAAAAAGATTTCTGAAAATAATGAGTTGC
AATGGTTGTCGAGTTCTACGAAAAGGTTGCAGTGAGAATTGCATCCTCCGTCCATGTATTCAATGG
ATCGAATCACCTGAAGCTCAAGGCCACGCCACCGTCTTCGTCGCTAAGTTCTTCGGCCGTGCCGGT
TTAATGTCTTTCATCTCCGCCGTACCGGAATCTCAATGCCCTGCTTTGTTTCAGTCTTTGCTATAC
GAAGCTTGTGGGAGAACTGTGAATCCGGTGAACGGAGCCGTCGGATTGTTGTGGACGGGGAATTGG
AATGTTTGTCAAGCGGCGGTTGAGACGGTGCTTCGTGGTGGTTCTTTAAAACCAATACCGGAGCTT
CTTAACGGCGGTGGATTCGCCGGGTTTCCGTCTCCTACTTCCGACGAAGCTTCGGAGATCTGTACG
GAAATGTTGAATCTACGAAAAGCTGATGATTCCGGTGATCGGAACATTTATCATCACTGCCGATTC
TCAAGCTCTAGATCTAGATCAAGATCAACAGCTTCTCCGCCGAAACGGAAACGATTATCGTCGGAA
CAACAACCTTCGTCGGAGCTTGATCTCTCTCTTATTCCTATTTATCCGATTAAAACCTTGCCGTTT
AAGGAAGATACACCGTCGATGTACTCGGAGGAGTCTGTTACCACGGTTTCGTTTCAAAACAACAAC
GCCGGTGATCGGTACGTACGCTGCGGCGGAGGAGGAGGAGGAGCAACGACAAAGTTGCTCAATCTC
TTCGCTTGAACGGCGTCGTTTTTGTTATGAGGGTTAAAAATGTAATTTCGAGTAACTTTTTGTTGT
GGCCTTTTTTGGATTAATAGCCCTTAATATTTTGTGAAATGTAGCAAATTACGTGATTTATTTCTC
TCTTTAATTATGAGAAATTATCCAAAGTAATAGAGAATAATGATAT

FIGURA 5 CONT.

SEQ ID NO: 65, At-LBD39-AY072144

GTTTTTTCTTTCCCTTCTTCAATCAAAACCTATTTGCATGCTCTCAAACCCGAATTAAATCGACAC
TTTTCAGTTTTTGTTTTAACAAGTAGAGTTTCCCAAAATATTGGATATATTTCTTTTTCAAATTTC
GGAAAAGAAATGAGTTGCAATGGATGTAGAGTTCTTCGAAAAGGTTGCAGTGAAACATGCATCCTT
CGTCCTTGCCTTCAATGGATCGAATCCGCCGAGTCACAAGGCCACGCCACCGTCTTCGTCGCTAAA
TTCTTTGGTCGTGCTGGTCTCATGTCTTTCATCTCCTCCGTACCTGAACTCCAACGTCCTGCTTTG
TTTCAGTCGTTGTTGTTTGAAGCGTGTGGGAGAACGGTGAATCCGGTTAACGGAGCGGTTGGTATG
TTGTGGACCAGGAACTGGCACGTATGCCAAGCGGCGGTTGAGACTGTTCTTCGCGGCGGAACTTTA
CGACCGATATCAGATCTTCTTGAATCTCCGTCGTTGATGATCTCCTGTGATGAGTCTTCAGAGATT
TGGCATCAAGACGTTTCAAGAAACCAAACCCACCATTGTCGCTTCTCCACCTCCAGATCCACGACG
GAGATGAAAGACTCTCTGGTTAACCGAAAACGATTGAAGTCCGATTCGGATCTTGATCTCCAAGTG
AACCACGGTTTAACCCTAACCGCTCCGGCTGTACCGGTTCCTTTTCTTCCTCCGTCGTCGTTTTGT
AAGGTGGTTAAGGGTGATCGTCCGGGAAGTCCATCGGAGGAATCTGTAACGACGTCGTGTTGGGAA
AATGGGATGAGAGGAGATAATAAACAAAAAAGAAACAAAGGAGAGAAAAAGTTATTGAACCTTTTT
GTTTAAAACCGACGACGCAAAACACTCAAAGATTTTGAGGCTCTCTTTTTTAGGGTTTTGAGTGGG
AATGGATATTTAGTTAATGATTTTTCTCTATCGAGAAATATGATAAAATTTTGGGGAAAAAAAAAAA
AAAAAA

SEQ ID NO: 66, At-LBD40-BT010943

CTCAAACCCAAAACACATTTTGTTTTTTGTTTCCCATCTAAAATGCGTATGAGTTGTAACGGATGT
CGAGTTCTACGAAAAGGCTGCAGTGAAAACTGTAGCATAAGACCTTGTCTTCAATGGATCAAATCC
GCTGAATCTCAAGCCAACGCCACCGTCTTCCTCGCCAAGTTTTATGGCCGTGCGGGGCTCATGAAC
CTCCTCAACACCGGTCCTGACCACCTCCGTCCTGCGATATTTAGATCATTGTTGTACGAAGCATGC
GGGAGGATTGTGAATCCGATATACGGATCCGTGGGGTTATTGTGGTCAGGAAACTGGCATCTTTGT
CAAGCGGCCGTTGAGGCAGTAATGAGAGGCTCACCCGTGACTCCGATCGCATGCGACGCAGCGGTT
ACTGGTCAAGCACCTCCTTTTAACAACAAGCTTTGCGACATAAGACACGTGTCAAGCAGGGACGAG
AACGTTAAGAGACGGAGCCGTGGTGCATGCAAGGAAGAGAGAAACGTGAGGTCTTTGAGTCATGAG
TCGTCACTGAGTCACGAGTCACCGGTGTCTTCTGAGGAGACGACGACGGAGGAACCAAAGACTTGG
ATCGGGCTTGAGCTGACTTTGGGGTTGGAGCCTTTAGCACGTGGAAATCACGTGGTGGTACCGATG
AAGAAAAGAAAGTTAGAGAGGTGTGGCACGTCTGAGGATGAGGACACGTGTAAGATTGAGCTTGGA
CTGGTGTGCAGTGAGTGAATGGTTCTTTTTTTGTGGCTGGTCTTAATTACAAGTTTTGGTGTTGAG
TTTTAGGTGTAC

SEQ ID NO: 67, At-LBD41-AF447895

CTCTCTCCGCTTCCCCCAAAAATCCAGGGGCTTCTTACACGCACACAATCCTTTAAAGCTATCGTC
TTTCTTATATAATCAGTGAATCTTCATTAACACCTCCCACAAAATCTCAGAAATAACTTTCACAAC
AGAGTCAAAGAGTTCCAAATCGTTTGTTTGTATTTTGTACGATCAAAGTTGTTGGTACTTGTAAGA
TAATCGAAACCAAAGATGCGGATGAGCTGTAATGGATGCAGAGTTCTTCGGAAAGGGTGTAGTGAG
GATTGTAGTATAAGGCCGTGTTTGGCTTGGATCAAATCGCCTGAAGCGCAAGCTAACGCAACTGTC
TTTCTCGCCAAGTTCTATGGCCGTGCTGGACTCATGAACCTCATCAACGCCGGTCCCAATCACCTT
CGTCCTGGGATTTTCCGATCGTTGTTGCACGAAGCTTGTGGGAGGATTGTGAATCCGATCTATGGT
TCGGTGGGTTTGTTGTGGTCGGGGAATTGGCAGCTTTGTCAAGACGCCGTGGAGGCTGTGATGAAA
GGAGAACCGGTCAAAGAGATCGCCACAGACGCTGCGACGATCGGCCAAGGTCCGCCTCTTAAGATC
TACGACATCCGACATATCTCCAAGGATGATAACTCTGCCGCCGCGGCTACTGGCTCAACCGATTTG
AAACTTGCGAAAACTCGCCGTGCTAAGCGGGTCTCCACCGTCGCGATACAGGCGGAATCGGAGGGA
AAGTCTGACGAGGCTAGTCACGATTCGTCGTTGAGTCATCAGTCTGAAATAGTGGCTGCTCATGAA
GGAGAGAGCAAGGAATCCGAGAGCAATGTCTCTGAGGTTTTGGCATTCTCGCCTCCGGCTGTGAAG
GGCTCCGGCGAGATAAAGCTTGACCTAACTTTAAGGCTCGAACCGGTGTCACGTGCGTATCATGTG

FIGURA 5 CONT.

GTACCTGTTAAGAAGAGAAGGATCGGCGTGTTTGGCACGTGTCAGAAGGAGAGCACGTGTAAGACT
GAGCTTATGCTCTAAGCTCCGTTTTAGCTATACTCAGAAGTTTTGCCCCCATCAGCTCTATTCGGC
TCGGCTCTGATGATCTTCAAGTCGAGCCGGTAACTTTTTTTTAATATATAGTTCTCTATTTCCCCC
TTTAATGTTCACTTTCCATTTGTGACAAAAGAATAGAGTATATGTACAAAATTATTTACAGAGATT
AATTCAAGAATTGGCTGAGATATTATCTA

SEQ ID NO: 68, At-LBD42-AF447896

CAAAAGTTGAAAAATACCAGCTCATTCACTTATGACTCATGAGTAACTTCTTCCTTATAAATTCAA
ATTGTTCACAACTTCCTTGAGTTCCATTTCCGAACACAAACACACATCATACGTATATTTGTCACT
AATAAACATATAAAGGAATATGAGAATCAGCTGCAACGGGTGTAGGGTTCTACGCAAAGGTTGCAA
CCAAGATTGCACCATTCGACCATGCCTTCAATGGATCAAATCCGCAGACTCTCAAGCCAATGCTAC
ACTCTTCCTTGCTAAGTTCTACGGTCGAGCCGGTCTTCTTAACCTCATCGAATCCGGTCCTGATCA
CCTTCGTCCCGCAATATTTAGGTCACTTCTGTACGAGGCTTGTGGTCGGATCGTGAACCCTGTAGA
TGGTTCGGTTGGTTTGATGTGGTCAGGGAACTGGGCTCAGTGCCAAGCAGCTGTTGACGCCGTGCT
CAACGGCTTACCCATAACTCACACACCTCTTCCAAGTGCATCCGCGTCGCACCAGATCATTCCTCC
CCACAGGACTTACGATATACGCCACGTGGCTAAAGATCCAACAACCGGCGGCGACAGTTCGGAGAA
TCTGGCCACACGTGTAAACGCTAACAAGAGCAAGACGCAAACTGGCCGATTCAAACGTGAATCCGT
GAACCAACTAGGTGAATGTAGTCATGACATGTGGCAACTCCCAAGTTCTAGTGCTACGCATGGTTA
TGGCCACTTCACGTTGGAGAATGTGGAGAGCCGAAGGGAAGCTCCATTTAATCAAAGTTCTCCAAA
TCTTGGGTTTGATGATCAAGTCGATATCAACGAGGTTGGCTTAGAGCTCAGACTTGGTTAGAACCA
GAAGATTGCATTATTATTATCATAATTCTGTCTTTTTAATTAATAATTTCACTTTCTATATAAAAG
GAGAAAAAAAAAACTAAAGCAAATCAAATTTAGTATTATTGTCAATCATTAGGTTATCAT

SEQ ID NO: 69, Linum usitatissimum

GTCTCCAACCCCCAAACACTCAAAAAATTATAGAAAATTACCAGCCGCCGGAAGCAGAGGTGATCG
GAGGATGAGCTGCAACGGTTGCAGAGTACTGAGGAAAGGATGTGGAGAGAATTGCGTGCTGAGGTC
ATCCCTCCGATGGATTCCCACTCCTCAAGCACAAGGCAACCCACTTTGTTCCTCGCCAAATTCTTC
GGCCGCAGCGACCTCCTCTCCTTTATCTCCTCCGTCCCCGATCCTCAACGCCCCGCTCTGTTCCAG
TCGCTGCTATTTGAAGCGTGTGGGAGGACGGTGAACCCGGTTAACGGGGCGGTGGGGCTGCTGTGG
AGCCGCAATTGGCACGTGTGTCAGGCTGCGGTCGAGACCGTTCTCGCCGGAGGATCACTGCATCCG
CTTCCGGGAATTGCGGCAGGGGTTGCGCCGCCGAATTTCGACGAGTCGGATGACAGCTGCTCCGCC
GCCGGATTCGCTGCTAATCAGCTCCGCCCGGCGATTTTGGCGATTACTGGCGGGAATCGGAATCTA
TCATCGTCCTCCTCCACCTCGGAGTATCGACGTCGTCATAATGCTGCTCTGGTCGGTGGTGGGAAT
GAGAGGAGGATTAATAACAGAGATGCGAGGGCGTCGTTTTTTGCGGAGGAATCGGAAACGACGACA
TTCGAGAGCACCGGCGGCGATCGGAAGAAGCTGCTGAATCTTTTTGTTTGAGTAGAGAGGAGGAAA
GAAAAAAGAGAAGCCGAGTTTTTGTCGGCAAAGGAAAGAACAAGACTTGAATGCGGTTACTGGAAT
GTGAACTGTACTTTGGGTTTCTGTGTTTTGTTTCTTATGAACTTTATATGATAATATATATACATA
AGAGAGAGAAAGAAAGAACGCCGGAATGACAGTCCGGCGAGTTTTGGGGGAAAAAAAAAAAAAAAA
A

SEQ ID NO: 70, Linum usitatissimum

MWRELRAEVIPPMDSHSSSTRQPTLFLAKFFGRSDLLSFISSVPDPQRPALFQSLLFEACGRTVNP
VNGAVGLLWSRNWHVCQAAVETVLAGGSLHPLPGIAAGVAPPNFDESDDSCSAAGFAANQLRPAIL
AITGGNRNLSSSSSTSEYRRRHNAALVGGGNERRINNRDARASFFAEESETTTFESTGGDRKKLLN
LFV

FIGURA 5 CONT.

SEQ ID NO: 71, Triticum aestivum - TC235711

ATGCGGCTTAGCTGCAACGGCTGCCGCGTGCTGCGGAAGGGCTGCAGCGAGGACTGCAGCATCCGC
CCCTGCCTGCAGTGGATCAAGAGCCCCGAGGCGCAGGCCAACGCCACCGTCTTCCTCGCCAAGTTC
TACGGCCGCGCGGGGCTCATGAACCTCATCAACGCCGGCACGGACGACAGCCTCCGCCCCGGCATC
TTCCGCTCGCTCCTCTACGAGGCCTGCGGCCGGATCGTCAACCCCATCTACGGCTCCGTCGGCCTG
CTCTGGTCCAACAACTGGCAGATGTGCCAGGCCGCCGTCGAGGCCGTCCTCAGCGGCAAGCCCATC
GTCCAGGTCTCCTCCGAGGATGCCGCCGCCGACCGGACACCGCCGCTCAAGGCGTACGACATCCGC
CACGTCTCCACCTCGCCAGCCGCCGACGGGAGGCTCCACAAGGTCGCCAAGCCGGGTCGCACCCGC
TTCAAGCGCGCGTCCTCCGCCTCGTCCCACCACAACCCGTCCAGCGACTCCAACAACAAGCCCAAG
CCCAAGCCGCAGCCTCAGACGCGGGCGCCCACGGCGGAGGAGGAGCGGGATCGTCAACACCGCAAG
GAGATGGAGGAGGGCGCGTTCCAGCGTGCTCCGAGCCACGAATCCTCCGACAGCCGACACGAGGAC
CCCGTGGAGCCACACTGCCAGCAAGAGGCGTCCGCGGACACGGAGGCCGAGGCGGGGTCGCACGTC
AGCCAGGCAGAGCAGGAGCAGAGCACAGAGCCGGCCGCAGAGCACGCGGAGGAGGTCGAAAAGGAG
GACGAAGAACTAGGGCTCGAGCTCACGCTTGGATTCGCTCCCGTCGCAGCACGGCCTGCCGAATTC
ACCTGACCGTCCGAACCGGGGGGAGCGGCTTTCGCG

SEQ ID NO: 72, Medicago truncatula - ABE82505

ATGAGTTGCAATGGGTGCCGAGTTCTTCGAAAAGGTTGCAGTGAGTCGTGTATTTTACGACCTTGT
TTGCAATGGATCGAAACTCCAGAAGCTCAAGGCCACGCTACTGTTTTCGTAGCAAAATTCTTTGGT
CGGGCTGGTCTCATGTCCTTCATTTCCAACGTCCCTGAACCACAGAGACCAGCTCTGTTTCAGTCT
CTACTGTTTGAAGCGTGTGGGAGAACGGTTAACCCTGTGAACGGTGCTGTTGGACTTCTTTGGACT
GGAAACTGGCACGTGTGTCAAGCAGCAGTTGAGACGGTTCTTCGCGGTGGCACGCTAAGGCCTTTA
CCGGAGCTGACACTTCTGGACGCGCCTCTCACTGCCACCGATGAAGCTTCTGAGGCGGAAGACGGT
GGTGCTGATGCTATGTGGAAGATCCGAGAACCGAATTTCGGGTTGGCGAGCTCACGATCTAGCAAT
AAGGTGTGTTCTGGAAGCAAACGGAGGAGATCGGAGGAGTTTGTCAAGGTTCCGGCGGCGATTAAT
CTTGATTTACGATTGACGCCGATTTTTCAACAGAAGGCGGTGGAGGAACGCCGGCATGGTTCGCCG
TCGATGACGTCGGAGGAATCGGTGACGACGACGGCGTGTTTGGAGACTGGGATTGGAGATCGGTGG
AGTCACGGTGGTGATCGGAAAGTTCTCAACCTTTTCATATGAGGATTTTCTGTATTTGTGTTGAGC
GTGACTCTCACGCGCCGTAAAGTTTTTCTGTTTTGCTCACAATACTCATTGTGGTGAAAAGAGTAA
AACGGTTTTGTAGTTCGGTGAATTTTCGTGAGTTACCGGTTCAGGATTTCTCTAGAGAGAAAGATG
ATGGTTCGTCGGTGATTTTTTTCTTTTTGCTGCTAAATATGAGATGTGTAAACCTGTAACCGTGGA
ATATCAAATATTAAATCAAATTAAATAAATAGTTTTTTTTTTTTTAATCATCGCTGTTTGAAATTT
AGTAGTAGTTGGATAAACATTCAGTTATTATAGTTCAATTTGAACTTCATATATTTTAATCATATA
AACTCAATATTA

FIGURA 5 CONT.

MSGATTASSGDHNNLRLPTPTLDAESQTLLQSISAEGGYAYARM
AVLAVAGDQSAAEAARDMAWEQLHSGPWHSVLPVWRDAYSMACL
HVAKIHFAAGEFGEALGALDMGLIMGGMLLRKDLHDSVLLVSS
EARKMTKSLEEASGDFKGERLVPEVPVDVNEVLKILPCRSLTC
KRVEKRSGLSLEGFLRDYYLPGTPVVITNSMAHWPARTKWNHLD
YLNAVAGNRTVPVEVGKNYLCSDWKQELVTFSKFLERMRTNKSS
PMEP**TYLAQHPLFDQINELRDDICIPDYCFVGGGELQSLNAWFG**
PAGTVTPLHHDPHHNILAQVVGKKYIRLYPSFLQDELYPYSETM
LCNSSQVDLDNIDETEFPKAMELEFMDCILEEGEMLYIPPKWW
HYVRSLTMSLSVSFWWSNEAESSSS

FIGURA 6

FIGURA 7

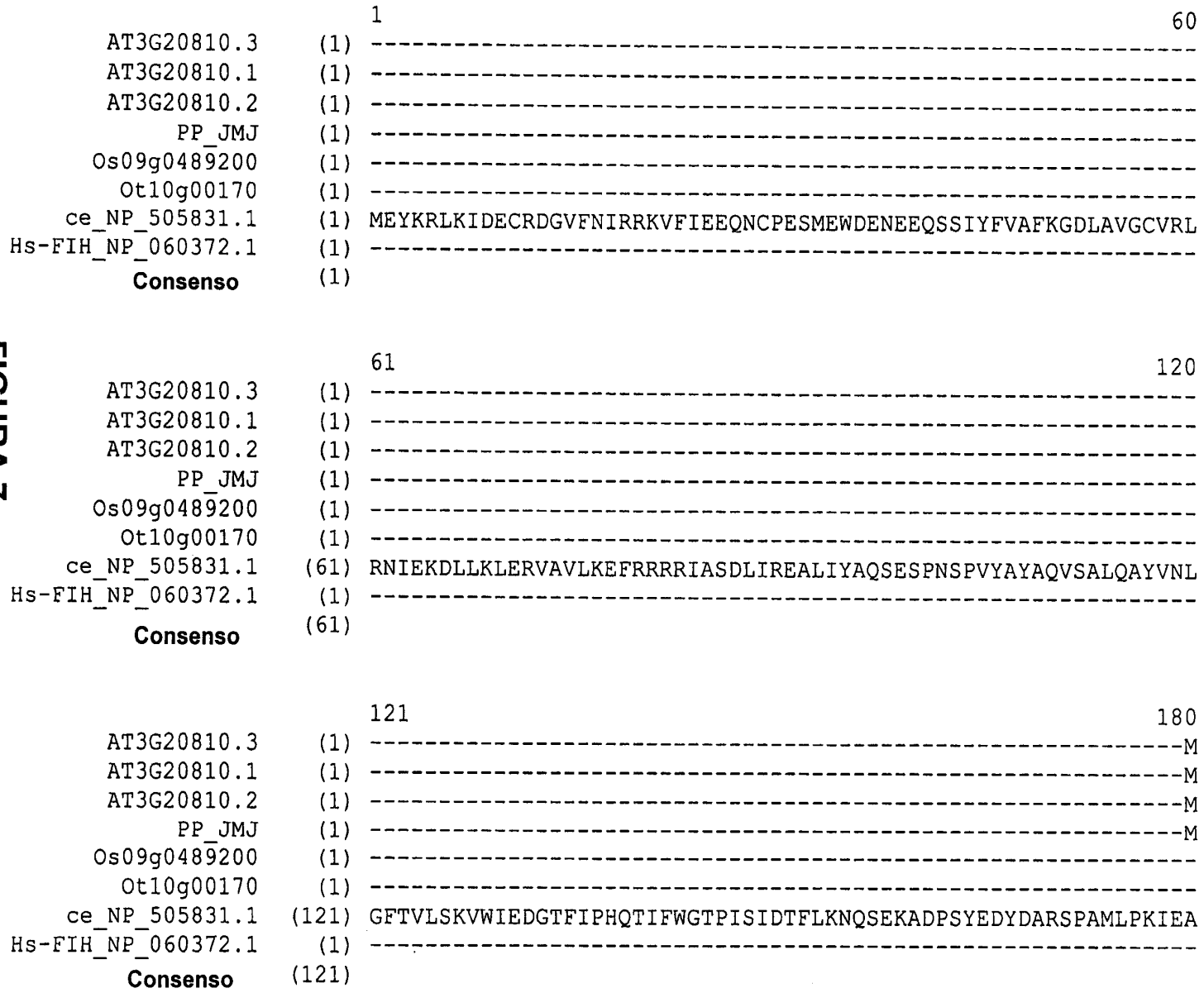
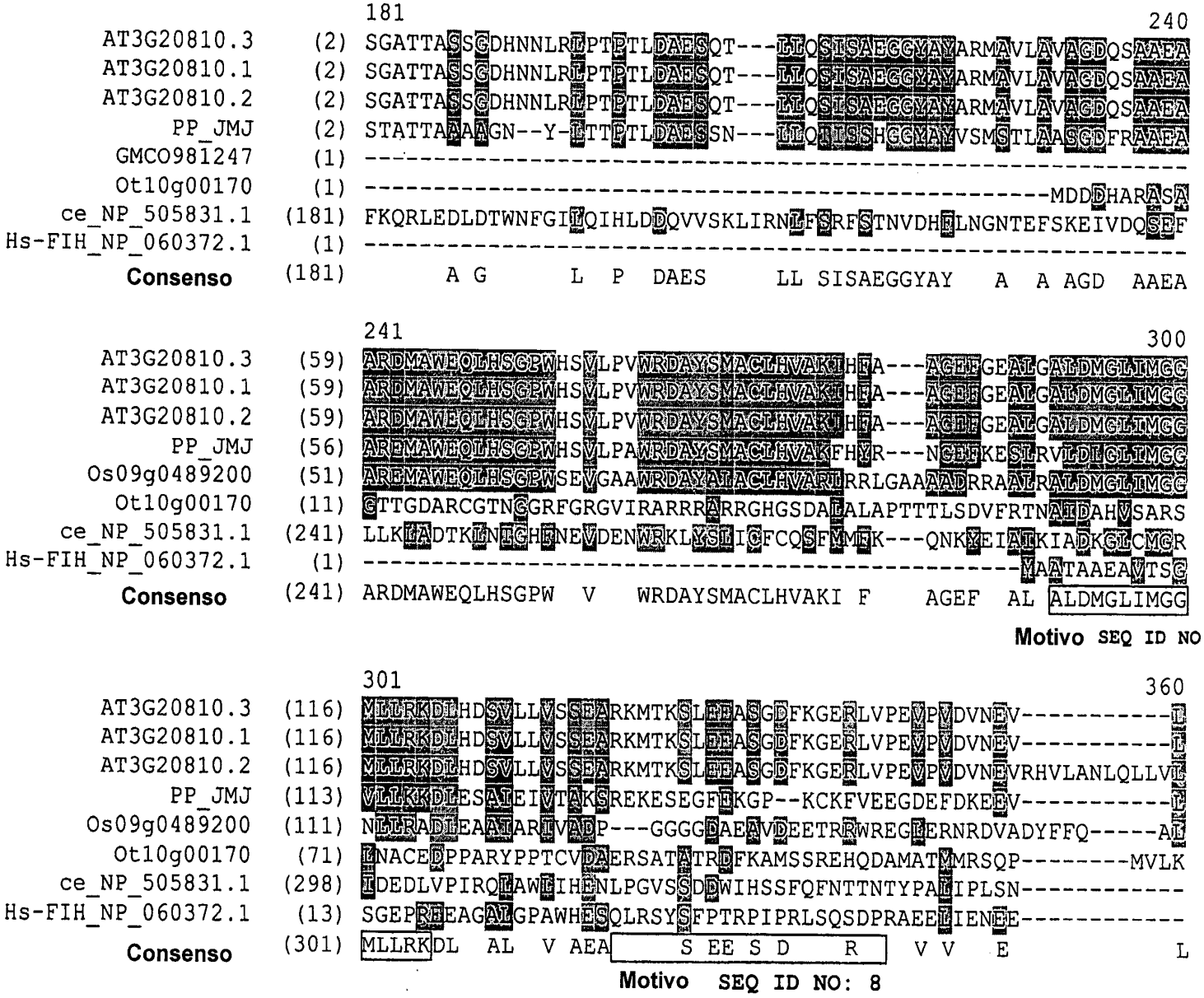

1 60
AT3G20810.3 (1) ----------
AT3G20810.1 (1) ----------
AT3G20810.2 (1) ----------
PP_JMJ (1) ----------
Os09g0489200 (1) ----------
Ot10g00170 (1) ----------
ce_NP_505831.1 (1) MEYKRLKIDECRDGVFNIRRKVFIEEQNCPESMEWDENEEQSSIYFVAFKGDLAVGCVRL
Hs-FIH_NP_060372.1 (1) ----------
Consenso (1)
61 120
AT3G20810.3 (1) ----------
AT3G20810.1 (1) ----------
AT3G20810.2 (1) ----------
PP_JMJ (1) ----------
Os09g0489200 (1) ----------
Ot10g00170 (1) ----------
ce_NP_505831.1 (61) RNIEKDLLKLERVAVLKEFRRRRIASDLIREALIYAQSESPNSPVYAYAQVSALQAYVNL
Hs-FIH_NP_060372.1 (1) ----------
Consenso (61)
121 180
AT3G20810.3 (1) ----------M
AT3G20810.1 (1) ----------M
AT3G20810.2 (1) ----------M
PP_JMJ (1) ----------M
Os09g0489200 (1) ----------
Ot10g00170 (1) ----------
ce_NP_505831.1 (121) GFTVLSKVWIEDGTFIPHQTIFWGTPISIDTFLKNQSEKADPSYEDYDARSPAMLPKIEA
Hs-FIH_NP_060372.1 (1) ----------
Consenso (121)

FIGURA 7 CONT.



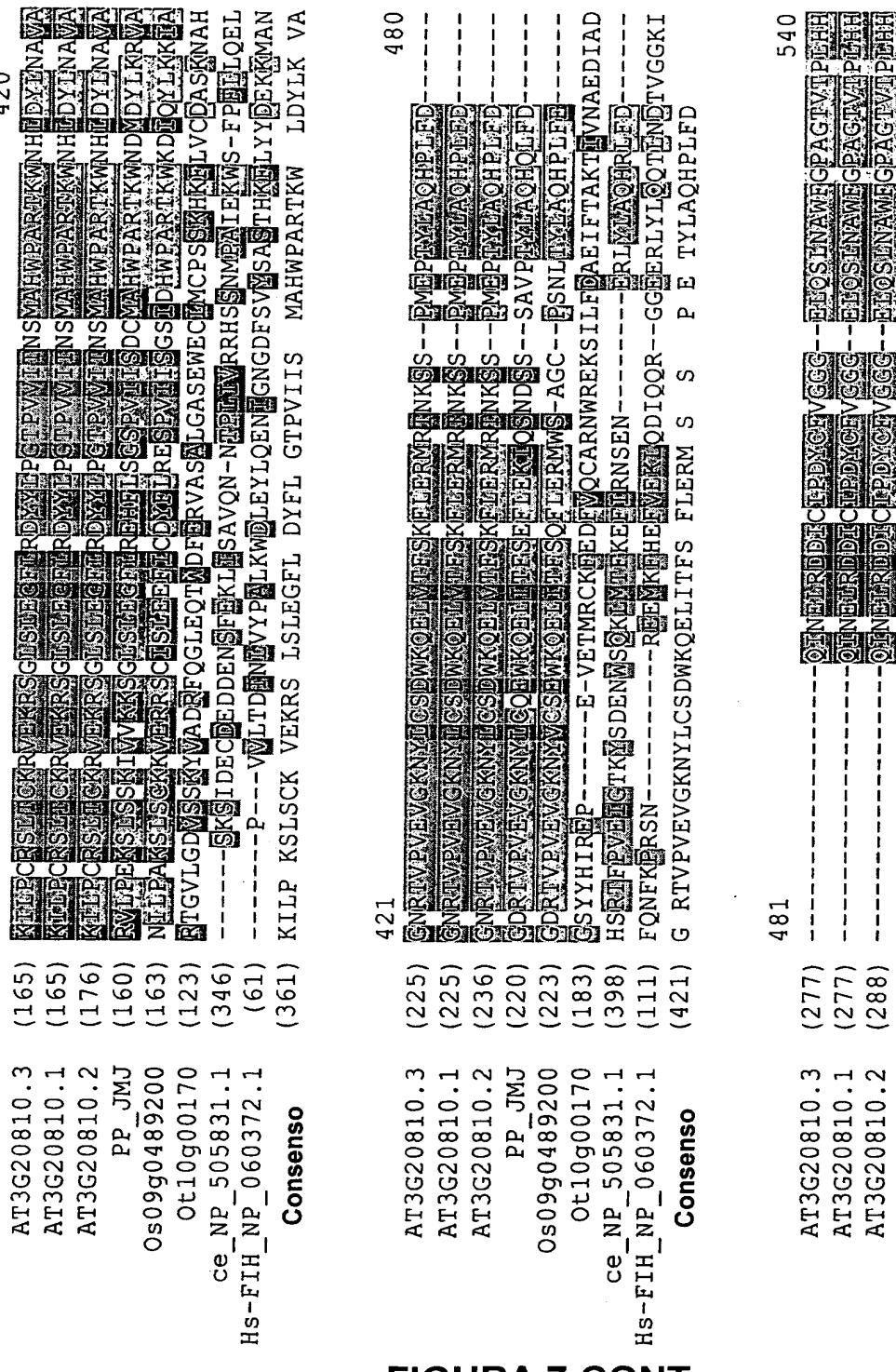


FIGURA 7 CONT.

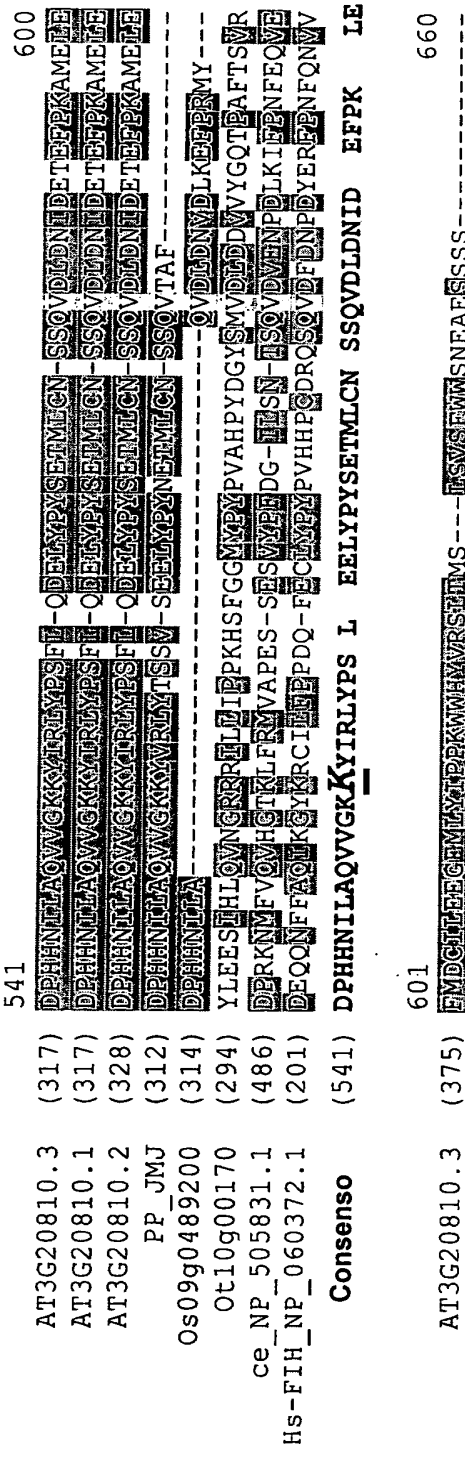


FIGURA 7 CONT.

```
                    721                                                780
AT3G20810.3  (419) ----------------------------------------------------
AT3G20810.1  (399) ----------------------------------------------------
AT3G20810.2  (430) ----------------------------------------------------
      PP_JMJ (356) ----------------------------------------------------
Os09g0489200 (368) ----------------------------------------------------
     Ot10g00170 (473) RTMIDGRLDPTPWLNISFREPLYLTDKPVHVADTRNELEKEFPEFYVAKLKKEGYAVDYT
   ce_NP_505831.1 (579) ----------------------------------------------------
Hs-FIH_NP_060372.1 (341) NTMIKGRYN-------------------------------------------------
          Consenso (721)

                    781                799
AT3G20810.3  (419) ----------------------
AT3G20810.1  (399) ----------------------
AT3G20810.2  (430) ----------------------
      PP_JMJ (356) ----------------------
Os09g0489200 (368) ----------------------
     Ot10g00170 (533) PVSVFNPTHPETIASVPKA
   ce_NP_505831.1 (579) ----------------------
Hs-FIH_NP_060372.1 (350) ----------------------
          Consenso (781)
```

FIGURA 7 CONT.

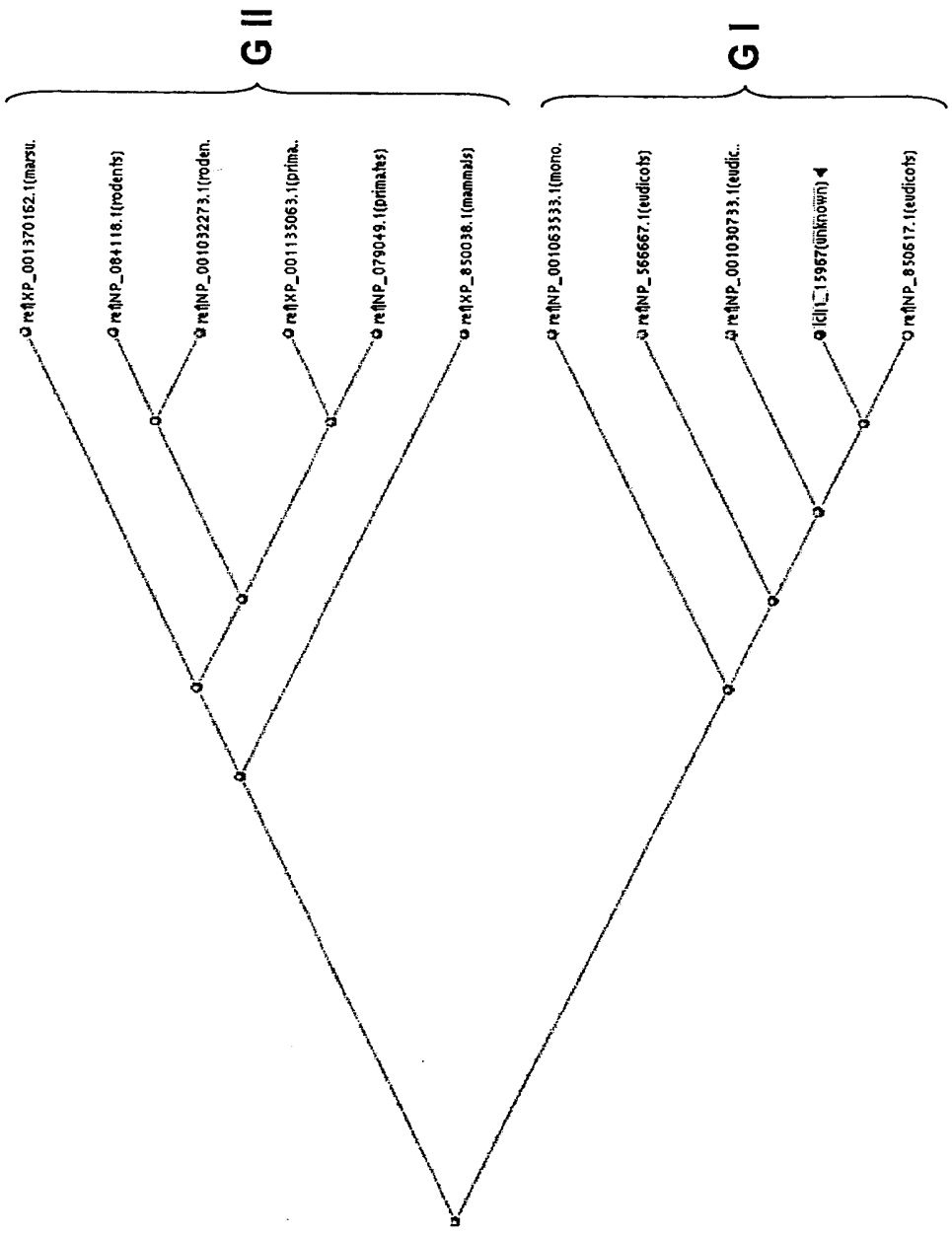

G II
G I
ref|XP_001370162.1|(marsu.
ref|NP_084118.1|(rodents)
ref|NP_001032273.1|(roden.
ref|XP_001135063.1|(prima..
ref|NP_079049.1|(primates)
ref|XP_850038.1|(mammals)
ref|NP_001063333.1|(mono.
ref|NP_566667.1|(eudicots)
ref|NP_001030733.1|(eudic..
ref|NP_850617.1|(eudicots)

FIGURA 8

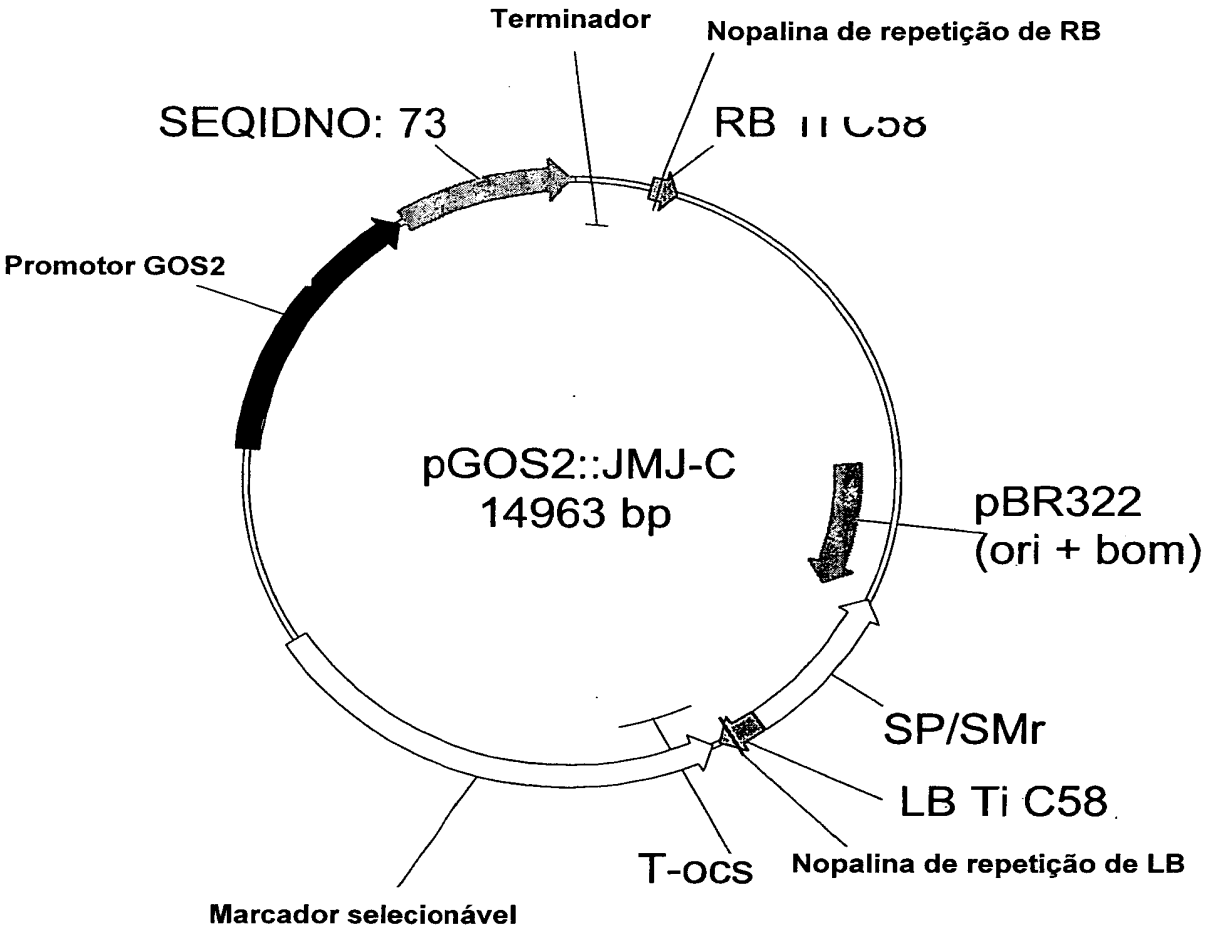

Terminador
Nopalina de repetição de RB
SEQIDNO: 73
RB Ti C58
Promotor GOS2
pGOS2::JMJ-C
14963 bp
pBR322
(ori + bom)
SP/SMr
LB Ti C58
T-OCS
Nopalina de repetição de LB
Marcador selecionável

FIGURA 9

SEQ ID NO: 73, DNA - *Arabidopsis thaliana*

ATGTCAGGAGCTACCACCGCTTCCTCCGGAGACCACAACAACCTCCGTCTCCCAACTCCGACGCTT
GACGCCGAATCGCAAACTCTGCTGCAATCGATTTCTGCTGAAGGGGGTTACGCTTATGCTCGCATG
GCGGTGTTAGCGGTGGCCGGTGACCAAAGCGCGGCGGAGGCGGCGCGTGATATGGCTTGGGAGCAA
CTTCACTCTGGTCCGTGGCACTCGGTTCTCCCCGTTTGGCGCGACGCTTACTCAATGGCTTGTCTC
CACGTAGCTAAAATCCATTTCGCCGCCGGCGAGTTTGGTGAGGCGCTTGGTGCGCTTGATATGGGT
CTTATCATGGGAGGTATGCTTCTCCGCAAGGATCTTCATGATTCTGTTTTGTTGGTCTCCTCTGAA
GCTCGTAAGATGACGAAGAGTCTTGAAGAAGCTTCTGGAGATTTCAAAGGTGAGAGATTGGTCCCG
GAAGTTCCCGTCGACGTCAATGAGGTCCTTAAGATTCTACCTTGCAGGTCTTTAACTTGTAAAAGA
GTAGAGAAGAGGTCTGGTTTATCTTTGGAGGGGTTTCTGCGTGATTATTATCTGCCAGGTACACCA
GTTGTTATAACCAATTCTATGGCTCATTGGCCAGCCAGGACCAAGTGGAATCACTTGGACTACCTC
AATGCTGTTGCTGGTAACCGCACCGTTCCCGTGGAGGTGGGGAAAAACTATTTGTGTTCTGATTGG
AAGCAAGAGCTTGTGACATTTTCTAAGTTCCTTGAACGGATGCGGACCAATAAATCGTCTCCAATG
GAGCCTACTTATCTTGCCCAGCATCCTTTGTTTGATCAGATAAATGAACTGAGAGATGATATATGT
ATTCCTGATTACTGTTTTGTCGGTGGTGGGGAACTCCAATCACTTAATGCATGGTTTGGCCCGGCT
GGGACAGTTACTCCGTTACACCATGATCCACATCATAATATACTTGCTCAGGTTGTTGGCAAGAAG
TATATAAGGCTTTACCCATCCTTCCTGCAAGACGAACTTTACCCTTACTCTGAGACAATGCTCTGC
AACTCTAGTCAGGTTGATCTAGACAATATCGACGAAACCGAGTTCCCAAAGGCCATGGAGTTGGAG
TTTATGGATTGTATTCTAGAAGAAGGTGAAATGTTGTACATCCCTCCCAAATGGTGGCACTATGTC
AGGTCCTTAACAATGAGTCTCTCGGTTAGCTTCTGGTGGAGCAATGAAGCAGAATCTTCTAGCTCG
TAG

SEQ ID NO: 74, proteína - *Arabidopsis thaliana*

MSGATTASSGDHNNLRLPTPTLDAESQTLLQSISAEGGYAYARMAVLAVAGDQSAAEAARDMAWEQ
LHSGPWHSVLPVWRDAYSMACLHVAKIHFAAGEFGEALGALDMGLIMGGMLLRKDLHDSVLLVSSE
ARKMTKSLEEASGDFKGERLVPEVPVDVNEVLKILPCRSLTCKRVEKRSGLSLEGFLRDYYLPGTP
VVITNSMAHWPARTKWNHLDYLNAVAGNRTVPVEVGKNYLCSDWKQELVTFSKFLERMRTNKSSPM
EPTYLAQHPLFDQINELRDDICIPDYCFVGGGELQSLNAWFGPAGTVTPLHHDPHHNILAQVVGKK
YIRLYPSFLQDELYPYSETMLCNSSQVDLDNIDETEFPKAMELEFMDCILEEGEMLYIPPKWWHYV
RSLTMSLSVSFWWSNEAESSSS

SEQ ID NO: 75, DNA - *Oryza sativa*

AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAATAT
AAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCACCTACT
TTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGTAATTAAGT
GGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGTTAAATTATTCG
AGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTTCTAGCTGAACTCAATGGGT
AAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTCTGAACGTATTGGCAAAGATTTAAA
CATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACATCATTAAGGACATGTCTTA
CTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATTTTTATTATTTATCTTTTTTCG
ATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAAT
ACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTAGGTAGCAATATC
TGAATTCAAGCACTCCACCATCACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTT
ACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAAAAAAGAATT
TTGCTCGTGCGCGAGCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCCACA
GAACAACCCACAAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAG
GCTTTGCGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGCATCCTCCTTCTCCCATCTATAA
ATTCCTCCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAAG

FIGURA 10

GACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTCGATCCATATCTTCCGGTCGAGTTCTTGGT
CGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCTCCCTTCGGTTGTTCTTGGATTT
ATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCC
TGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTCGGTTTGATTAGTAGT
ATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTAGGGATCGGAATCTTGCGATT
TTGTGAGTACCTTTTGTTTGAGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAGTAC
GGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTC
CCTATTGAACAAAAATAATCCAACTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTT
AAGCCTGTCCAAAATTTCGCAGCTGGCTTGTTTAGATACAGTAGTCCCCATCACGAAATTCATGGA
AACAGTTATAATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTT
TTTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATG
CTTTTATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGA
AGAACTTATCCGATTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATT
CATTTGGATTATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAA
CTGTCCTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGT
AGAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAATCG
GGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCTTGCCACT
TTCACCAGCAAAGTTC

SEQ ID NO: 76, DNA – sequência artificial – iniciador – 1
GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGTCAGGAGCTACCACCG

SEQ ID NO: 77, DNA – sequência artificial – iniciador – 2
GGGGACCACTTTGTACAAGAAAGCTGGGTAAAATAAATCCATTCGCTACG

SEQ ID NO: 78, Proteína - Sequencia artificial-c domínio de JMJ-C da SEQ ID N°: 74

TYLAQHPLFDQINELRDDICIPDYCFVGGGELQSLNAWFGPAGTVTPLHHDPHHNILAQVVGKKYI
RLYPSFLQDELYPYSETMLCNSSQVDLDNIDETEFPKAMELEFMDCILEEGEMLYIPPKWWHYVRS
LTMSLSVSFWWSNEAESSSS

SEQ ID NO: 79, Proteína - Sequencia artificial-caixa conservada em domínio JMJ-C da SEQ ID N°: 74

CILEEGEMLYIPPKWWHYVRSLTMSLSVSFWW

SEQ ID NO: 80, Proteína - Sequencia artificial - Domínio DUF 335
RKMTKSLEEASGDFKGERLVPEV

SEQ ID NO: 81, Proteína - Sequencia artificial - motivo conservado
ALDMGLIMGGXLLRKDL

SEQ ID NO: 82, Proteína - Sequencia artificial -sequência de consenso de ligação de Fe(II)

HX(D/V)

FIGURA 10 CONT.

SEQ ID NO: 83, DNA - *Arabidopsis thaliana*

ATGTCAGGAGCTACCACCGCTTCCTCCGGAGACCACAACAACCTCCGTCTCCCAACTCCGACGCTT
GACGCCGAATCGCAAACTCTGCTGCAATCGATTTCTGCTGAAGGGGGTTACGCTTATGCTCGCATG
GCGGTGTTAGCGGTGGCCGGTGACCAAAGCGCGGCGGAGGCGGCGCGTGATATGGCTTGGGAGCAA
CTTCACTCTGGTCCGTGGCACTCGGTTCTCCCCGTTTGGCGCGACGCTTACTCAATGGCTTGTCTC
CACGTAGCTAAAATCCATTTCGCCGCCGGCGAGTTTGGTGAGGCGCTTGGTGCGCTTGATATGGGT
CTTATCATGGGAGGTATGCTTCTCCGCAAGGATCTTCATGATTCTGTTTTGTTGGTCTCCTCTGAA
GCTCGTAAGATGACGAAGAGTCTTGAAGAAGCTTCTGGAGATTTCAAAGGTGAGAGATTGGTCCCG
GAAGTTCCCGTCGACGTCAATGAGGTGAGACATGTTTTGGCTAATCTACAATTACTTGTCCTTAAG
ATTCTACCTTGCAGGTCTTTAACTTGTAAAAGAGTAGAGAAGAGGTCTGGTTTATCTTTGGAGGGG
TTTCTGCGTGATTATTATCTGCCAGGTACACCAGTTGTTATAACCAATTCTATGGCTCATTGGCCA
GCCAGGACCAAGTGGAATCACTTGGACTACCTCAATGCTGTTGCTGGTAACCGCACCGTTCCCGTG
GAGGTGGGGAAAAACTATTTGTGTTCTGATTGGAAGCAAGAGCTTGTGACATTTTCTAAGTTCCTT
GAACGGATGCGGACCAATAAATCGTCTCCAATGGAGCCTACTTATCTTGCCCAGCATCCTTTGTTT
GATCAGATAAATGAACTGAGAGATGATATATGTATTCCTGATTACTGTTTTGTCGGTGGTGGGGAA
CTCCAATCACTTAATGCATGGTTTGGCCCGGCTGGGACAGTTACTCCGTTACACCATGATCCACAT
CATAATATACTTGCTCAGGTTGTTGGCAAGAAGTATATAAGGCTTTACCCATCCTTCCTGCAAGAC
GAACTTTACCCTTACTCTGAGACAATGCTCTGCAACTCTAGTCAGGTTGATCTAGACAATATCGAC
GAAACCGAGTTCCCAAAGGCCATGGAGTTGGAGTTTATGGATTGTATTCTAGAAGAAGGTGAAATG
TTGTACATCCCTCCCAAATGGTGGCACTATGTCAGGTCCTTAACAATGAGTCTCTCGGTTAGCTTC
TGGTGGAGCAATGAAGCAGAATCTTCTAGCTCGTAG

SEQ ID NO: 84, Proteína - *Arabidopsis thaliana*

MSGATTASSGDHNNLRLPTPTLDAESQTLLQSISAEGGYAYARMAVLAVAGDQSAAEAARDMAWEQ
LHSGPWHSVLPVWRDAYSMACLHVAKIHFAAGEFGEALGALDMGLIMGGMLLRKDLHDSVLLVSSE
ARKMTKSLEEASGDFKGERLVPEVPVDVNEVRHVLANLQLLVLKILPCRSLTCKRVEKRSGLSLEG
FLRDYYLPGTPVVITNSMAHWPARTKWNHLDYLNAVAGNRTVPVEVGKNYLCSDWKQELVTFSKFL
ERMRTNKSSPMEPTYLAQHPLFDQINELRDDICIPDYCFVGGGELQSLNAWFGPAGTVTPLHHDPH
HNILAQVVGKKYIRLYPSFLQDELYPYSETMLCNSSQVDLDNIDETEFPKAMELEFMDCILEEGEM
LYIPPKWWHYVRSLTMSLSVSFWWSNEAESSSS

SEQ ID NO: 85, DNA - *Arabidopsis thaliana*

ATGCCGACGGGATGCGAGTTTCCGATGATCAAAACCTTCGAGAATGCTCTTACTGCCGCCGACTTC
GAGTCTACCGTCGAGCTCACCAACTTCCCAGCAGTTTTTCGGGGATGTGCTAGTGTCTGGGATGCG
TATTCTAAGTGGAATCCATTCAACAGTGGCCTCGATTATCTGGAGGAACGTGCAGGTTCAGTTGAA
GTGGAGGCAATGCTTTCGAGGACGGCTCCAGTTTTCAATGGTGATATCAGAAGCCATGAGAGGGTG
TCATTGCCTTTCTCGGATTTTATCAGATTCTGCAAGCAGCATATGAGAGGCAAAGGGAATGGTTCT
GGTGTTGATGCTAAGTCTGCAGATTTAAATCCTATGTGTGAGGACTATAGACCTGGACAAATATAC
CTAGCACAGTTTCCAATTCTGAATGATGAGAAAGAAGAGAAGGTTCTACTTAAGATTCTGAGACAA
GATATCCAGACGCCCACATTTTTAGACGCAAAGTCACTTTCTTCTATCAACTTTTGGATGAATTCT
GCAGAAGCTCGATCAAGTACCCACTATGATCCACATCATAACCTTCTATGCGTAGTCTCAGGGCGT
AAAAAAGTTGTATTGTGGCCCCCTTCCGCCAGTCCCTCGCTCTACCCGATGCCTATATACGGAGAA
GCATCAAACCATAGTTCGGTTGGTCTAGAAAATCCAAATTTATCAGATTATCCAAGAGCAGAGCAT
TCACTGAAGCAGTCACAGGAGATTACTCTTAACGCTGGTGATGCAGTATTCATTCCTGAAGGTTGG
TTCCATCAGGTGGATAGTGATGAACTGACTGTTGCTGTCAACTTTTGGTGGCAATCAAACTATATG
TCTAACATGCCTGAACACATGGATTCGTATTATCTGCGTCGAATCACAAGAAGATTGATAGACAGA
GAAATGAGCCTGTTAGTCTCGAAGCCTTCTTCAACTGATTTAAGGCACCTGTCTGAGCATATTGAC
CAATCTCGCATTGAGATGGCAGAAGGTGGGAATGATAATATAGGCAATGAAAGCATTAAAAAGGGC

FIGURA 10 CONT.

CTCAGCACATTGCATGAAAAAGCGTCGCTGCACGATCTAGATCCTTCTGCTTCCCAGGCGCTTCAT
GATCTTATTTCTTTAGTCCATGACCACGTTAATGCTGTTGATACAAAAGATGATCGCGTTGCCCAT
TTGCTTTGGAATCTTGAAGCCTCTAGACTTCGGGATGTTCTCCTTGCAATGGCACGTTATTTCCCA
AGAACTTTAGAAGCATTAATACTCCACATGCTCTCACCCATTGCTGCGGAAGTTCTAACACAGAAG
TTTGATGAGATTGACCAACAGACTGGTGAAGAAGATAGGACCCAGTTCTTTCGGGAATTTTATAGT
GCATTCGATGATGAAGCAGCTGCGATGGATATAATTTTAAGTAGGAAAGAGGCTTTTGCATTTCAG
GCATTTAAGAATGTACTAGACAAGTTTTTGGGGGTCAACATTGCTTCACCAACAACAAACATTTGA

SEQ ID NO: 86, Proteína - *Arabidopsis thaliana*

MPTGCEFPMIKTFENALTAADFESTVELTNFPAVFRGCASVWDAYSKWNPFNSGLDYLEERAGSVE
VEAMLSRTAPVFNGDIRSHERVSLPFSDFIRFCKQHMRGKGNGSGVDAKSADLNPMCEDYRPGQIY
LAQFPILNDEKEEKVLLKILRQDIQTPTFLDAKSLSSINFWMNSAEARSSTHYDPHHNLLCVVSGR
KKVVLWPPSASPSLYPMPIYGEASNHSSVGLENPNLSDYPRAEHSLKQSQEITLNAGDAVFIPEGW
FHQVDSDELTVAVNFWWQSNYMSNMPEHMDSYYLRRITRRLIDREMSLLVSKPSSTDLRHLSEHID
QSRIEMAEGGNDNIGNESIKKGLSTLHEKASLHDLDPSASQALHDLISLVHDHVNAVDTKDDRVAH
LLWNLEASRLRDVLLAMARYFPRTLEALILHMLSPIAAEVLTQKFDEIDQQTGEEDRTQFFREFYS
AFDDEAAAMDIILSRKEAFAFQAFKNVLDKFLGVNIASPTTNI

SEQ ID NO: 87, DNA - *Arabidopsis thaliana*

ATGGCTAAAGAGATAGAGAATTTATGGAGAGAAGTGAGAGAATTGAGCTTAGGAACTAAAATCGAT
CGTTTTGATTCTCAACCGTCTCCGGTGAAGTTTCTCCGAAACTACGTGTCTCAGAGCAAGCCTTGT
GTAATCTCCAAAGCTATTACTCATTGGCCGGCGTTAAAACTCTGGTCAGATCCGGCTTACCTCACC
GGAGCTTTGTCAGATGACGTTGTCTCGCTACATCTCACTCCTAACGGCTGCGCAGACGCCGTCACC
GGAGATAGTGATCTTTGTTTCGCTTCGGCTCATGTTGAGAAAGTTCTGTTCCCGGAGGCTCTTAAG
GTTGTTCAATCGTCGTGTAAAGGACTGAAGGTTGGGTATTTGCAACAGCAAAACGATTGTTTTAGA
ACGGAGTATTCGACGGTTGCGTTAGACTGCGACGGTGATATTGAGTGGGCTACGGAGGCATTTGGT
TGTTCGCCGGAAGCTGTGAACCTTTGGATTGGTACTGATGACTCCGTGACTTCATTTCATAAAGAT
CACTATGAGAATCTTTACGCTGTTGTTTCAGGGGAGAAACATTTCCTTCTTCTTCCTCCGACTGAT
GTTCACCGTCTCTATATCGAACAGTATCCGGCAGCTAATTACTCTTACCACCGAGATACTGATGCG
TTCAAGCTTGAGGTTGAGGAGCCGGTTAGGCATGTACCTTGGTCTAGTGTTGATCCATACCCTTCG
CCAGAGAAGGAAGCAAGCGAGAGATTGAAATTTCCGTTGTTCTTTGATGGCCCGAAGCCGTTTCAT
TGTACTGTCAAGGCTGGCGAGGTTCTTTACCTGCCAAGCATGTGGTTTCACCATGTTAGTCAAACC
CCAGGAGATGGAGGTTATACCATTGCAGTGAACTATTGGTATGACATGCAGTTTGACATCAAGTAT
GCTTATTTTAACTTCTTGCAATCTTTATTATACAAGTCATCTTCACTCAATCCAGTTCTTTCTTGG
AGAGAAGACGAAGATTCAGAATCCAGTGATGCAGAGATTGCTCCCTGA

SEQ ID NO: 88, Proteína - *Arabidopsis thaliana*

MAKEIENLWREVRELSLGTKIDRFDSQPSPVKFLRNYVSQSKPCVISKAITHWPALKLWSDPAYLT
GALSDDVVSLHLTPNGCADAVTGDSDLCFASAHVEKVLFPEALKVVQSSCKGLKVGYLQQQNDCFR
TEYSTVALDCDGDIEWATEAFGCSPEAVNLWIGTDDSVTSFHKDHYENLYAVVSGEKHFLLLPPTD
VHRLYIEQYPAANYSYHRDTDAFKLEVEEPVRHVPWSSVDPYPSPEKEASERLKFPLFFDGPKPFH
CTVKAGEVLYLPSMWFHHVSQTPGDGGYTIAVNYWYDMQFDIKYAYFNFLQSLLYKSSSLNPVLSW
REDEDSESSDAEGSKFTPSATNLYLSSTYVSDGARFWKRVHVNSIKLGFESNGLIQSGLIDMYRKY
RLVINAEKVFRAMEGQTVLTVGGYGFEYYAEWFSH

FIGURA 10 CONT.

SEQ ID NO: 89, DNA - *Medicago truncatula*

ATGAATGAGAAGATAGAAGAACTATGGAGAGAAGTACGAGAATTAAGTCTAGGAAGCAAAAGAAGC
ATAGAACGTTTAGAATCAGCACCAACACCACTCCAATTCCATAGAAACTTCATCACTCCAAACAAA
CCTTGCATCATCTCCAACTCCATCTCCCATTGGCCCTCCCTCTCTCTCTGGTCCCACCCTTCTTAC
CTCACTCAATCCCTCTCTTCCACCACCGTCTCCCTCCATCTCACCCCCACCGGCTCCGCCGACTCT
TTAACCCCTCTCCCCTCTTCCCCCTCCTCCCTCTGCTTCGCCTCTGCCCATGTCCAAAACCTCCCT
TTCCCCGAAGCCCTCCGTCTCATCAACTCCTCTAACCCTTCCCAATGTGTCGCCTACGCGCAGCAA
CAGAACGATTGTTTTCGTTCAGAGTATGATTCCATTGTTAAAGATTGTGATCAACATATTGCTTGG
GCTACTGAAGCATTTGGTTTAGAGCCTGAAGCTGTTAATCTTTGGATTGGAAACAAACATTCATCT
ACTTGGTTTCATAAGGATCATTATGAGAATCTTTATGCTGTTGTTACTGGTCAAAAACACTTTCTT
TTGTTTCCTCCTACTGATGTTCATCGCTTTTATATTCGGAATTACCCTGCTGCTACTTATAAATAT
TACATGGAAACTGGAGAGTTTGATTTGGAACTTGACAAGCCAACTAGGTATGTGCCTTGGTGTAGT
GTGAACCCTTTTCCTTCTCCGGAGAACTTGGAGGACGAGATTTCAAAGTTTCCTTTATACTTCAAT
GGCCCTCCACCGTTTGAGTGTACTGTCAAGGCTGGGGAGATTCTTTACTTGTATGATTCTCTTCTT
TTGCATTGCATCATGCCAAGCATGTGGTTTCATCATGTTAGACAAAGTGGGGATGATGGAGAATTA
ACTATTGCAGTAAATTATTGGTATGATATGCAATTTGACATTAAGTATGCTTATTTCAACTTTTTA
CAATCTATTGATTACCGATCTCCTACGAGTCCAATGATGCCTGAGAAATTGTGTGAGGAGATAGAT
TTTGGTCCTGATGACGATGAGACTAGATGAATTATCATTTCTCAGTCGGTGATATTCGGCAATGTG
GTCCTTCGATCGAATGCTGA

SEQ ID NO: 90, Proteína - *Medicago truncatula*

MNEKIEELWREVRELSLGSKRSIERLESAPTPLQFHRNFITPNKPCIISNSISHWPSLSLWSHPSY
LTQSLSSTTVSLHLTPTGSADSLTPLPSSPSSLCFASAHVQNLPFPEALRLINSSNPSQCVAYAQQ
QNDCFRSEYDSIVKDCDQHIAWATEAFGLEPEAVNLWIGNKHSSTWFHKDHYENLYAVVTGQKHFL
LFPPTDVHRFYIRNYPAATYKYYMETGEFDLELDKPTRYVPWCSVNPFPSPENLEDEISKFPLYFN
GPPPFECTVKAGEILYLYDSLLLHCIMPSMWFHHVRQSGDDGELTIAVNYWYDMQFDIKYAYFNFL
QSIDYRSPTSPMMPEKLCEEIDFGPDDDETR

SEQ ID NO: 91, DNA - *Brachypodium sylvaticum*

TCAATCGTTCTTCTCCTCAAGATCACCTTCCAAGGCAGCATCTTTGTTGTCCAGCGGAGAATTACT
GATCTCCAGTGACCTCAGGAAGTTAAAGTACGCATACTTGATGTCAAACTGCATGTCATACCAGTA
ATTCACAGCAATGGTGAGCCCGTTAGGCCCGGGGCTCTGGCTGACATGGTGAAACCACATGCTCGG
CAAGTAGAGCACCTCGCCAGCACGGACGGTGCACCGTATCGGCCTGGGCCCCTCGAAGTAGAGCGG
GAAGGATGAGACCTGCGCCGCCATCTCCTCCGGCGACGACGGGTTCGGGTCCACGCTGCTCCACGG
CACGATCCTCTCGGGCTCTTCCATCTCCAGCTTAAGCCCCGTCAGCTCCTCCTCTCCTTCGTTCTC
CGTGACGTAGCGGGCGGCGGGGTAGTCGCGGACGTAGAGGCGGTGGTGCTCGGTGGGGGGCAAGAG
GAGGAAATGCTTCTCGCCGGAGAGGACGGCGTAGACGTTGTCGTAGTGGTCCTTGTGGAAGGAGGT
GACGGAGCAGGAGTTGCCGATCCAGAGGTTGACGGCCTCCGGGAGGCAGCCGAGCGCCTCGCTGGC
CCAGGGCACGTGCGCGTCCACGTCGCCGGCCACCGCCGCGTACTCCCCGCGCAGGCAGTCGTCCTG
CTGCTGCGCGTACGCCACCAGGCCGCCAGCGGCCGCCCGGTCGGAGCCCCTGATGAGCCGCACGGC
GGTAGGGAAGTCGACCCGGCGGACGTGCGCGGACGCGAAGCACCTGGCGCCGGGGAGGCACGGGTG
AGGGGCGAGGGCGTCGGCGCGGCCGTCGGGGGTGAGGTGGACGGAGACGTCGGTGGAGCGGAGCGC
GTCGGTGAGGTAGGATTCGGTGGGCCAGAGGGAGGCGGCGGGCCAGTGGCGGGTGGCGGCGGCGGA
GACGAGGAGCGGGCGGCCCGGGGAGACGTGGTCGCGGAGGAAGGCGAGCGGGGTCGGGGGCAGGTC
GGCGCGCGGCACGGCCTCCGGAGAGTGGAGACCCAGCAGGTCCCGTGATTCCGCCCACAGCTCCCG
CACCGCGCGCTCCAT

FIGURA 10 CONT.

SEQ ID NO: 92, Proteína - *Brachypodium sylvaticum*

MERAVRELWAESRDLLGLHSPEAVPRADLPPTPLAFLRDHVSPGRPLLVSAAATRHWPAASLWPTE
SYLTDALRSTDVSVHLTPDGRADALAPHPCLPGARCFASAHVRRVDFPTAVRLIRGSDRAAAGGLV
AYAQQQDDCLRGEYAAVAGDVDAHVPWASEALGCLPEAVNLWIGNSCSVTSFHKDHYDNVYAVLSG
EKHFLLLPPTEHHRLYVRDYPAARYVTENEGEEELTGLKLEMEEPERIVPWSSVDPNPSSPEEMAA
QVSSFPLYFEGPRPIRCTVRAGEVLYLPSMWFHHVSQSPGPNGLTIAVNYWYDMQFDIKYAYFNFL
RSLEISNSPLDNKDAALEGDLEEKND

SEQ ID NO: 93, DNA - *Oryza sativa*

ATGGAGCGCGCAGTGCGGGAGCTGTGGGCGGAGTCGCGGGACCTGCTGGGCCTCCACTCCCCGGAC
GACGCCGCCGCCGCCGACGCCGCGATGCCCCGCGCCGAGATGCCCCCGACGCCGCTCGCGTTTCTC
CGCGACCACGTCTCGCCGGGGCGCCCTCTCCTCGTGTCCTCCGCCGCCACCAGCCACTGGCCGGCC
GCCTCGCTGTGGCCGACCGACTCCTACCTCACCGACGCGCTCCGCTCCACCGCCGTCTCGCTCCAC
CTCACCCCCGACGGCCGCGCCGACGCCCTCGCCCCGCACCCGCGCCCGAGCCACCCCGGCGCCAAG
TGCTTCGCCTCCGCGCACGTCCGCCAGGTCGACTTCCCCACCGCCGTGCGCCTCATCCGGAGCTCC
GATCCGGCCTCCGGCCTGGTGGCCTACGCGCAGCAGCAGGACGACTGCCTGCGCGGGGAGTACGCC
GCCGTCGCCGGCGACGTGGACGCGCACGTGCCCTGGGCCAGCGACGCGCTCGGCTGCCTCCCCGAG
GCCGTCAACCTCTGGATCGGCAGCGCCTGCTCCCAGACCTCCTTCCACAAGGACCACTACGACAAC
ATCTACGTCGTCGTCTCCGGCGAGAAGCACTTCCTCCTCTTGCCCCCCACCGAGCACCACCGCCTC
TACGTCCGCGACTACCCCGCCGCCCACTACGCCGCGGAAGATGAAGCGGAGCTGAGGCTTAAGCTG
GAGCTGGAGGAGCCCGAGAGGATCGTGCCATGGAGCAGCGTTGACCCCTACCCGCCGTCGCCGGAG
GAGGCGGCCGCGCAGGCGTCATCCTTCCCGCTCTACTTCGAGGGGCCGAGGCCGATCCGCTGCACG
GTGCGCGCCGGCGAGATGCTCTACTTGCCGAGCATGTGGTTTCACCATGTGAGCCAGAGCCCCGGG
CCGAACGGGCTCACCATTGCAGTGAACTACTGGTATGACATGCAGTTTGACATTAAGTATGCCTAC
TTCAACTTCTTGAGGTCATTGGAGATCGATGGTAGTTCGTCGAAAAAGACGGATGCTTTGGAAGAC
GATCTCGAGGAGACGAATGATTGA

SEQ ID NO: 94, Proteína - *Oryza sativa*

MERAVRELWAESRDLLGLHSPDDAAAADAAMPRAEMPPTPLAFLRDHVSPGRPLLVSSAATSHWPA
ASLWPTDSYLTDALRSTAVSLHLTPDGRADALAPHPRPSHPGAKCFASAHVRQVDFPTAVRLIRSS
DPASGLVAYAQQQDDCLRGEYAAVAGDVDAHVPWASDALGCLPEAVNLWIGSACSQTSFHKDHYDN
IYVVVSGEKHFLLLPPTEHHRLYVRDYPAAHYAAEDEAELRLKLELEEPERIVPWSSVDPYPPSPE
EAAAQASSFPLYFEGPRPIRCTVRAGEMLYLPSMWFHHVSQSPGPNGLTIAVNYWYDMQFDIKYAY
FNFLRSLEIDGSSSKKTDALEDDL

SEQ ID NO: 95, DNA - *Arabidopsis thaliana*

ATGGGGACAGAGCTAATGAGAATCTGTGTGAAAGAAGATAGTGATGATTTGCCATCTGTTCCTCCT
GGTTTTGAATCATATGCAACTTTTACCTTAAAAAGAGTAGTCCCTGCTACTACTTCTGATAAAGCC
AAGACTCCTGCAATTGAAAGTGTTTCTGCAACCGAACAAGCTAAGATGGAAGTTGAATCTGATGAA
GCCAAAGCTGCTCGGGCTCTACGCCGTAGACCTTGGATTAATCATAGCGGATGTGATGATGATGGT
GATTGTGCTGCAAACAATGACAATGCTGCATCTCAAAATCCTGATCAAAATTGTGATGTGAAGCCA
GCTCTTCCAAAGGGTGTGGTCAGGGGATGTGAAGAATGTAAAGATTGTCAGAAGGAGTTTGAAGAC
ACTTTAAATTATATCGCTAAAATAAGACCGGAGGCTGAAAAGTATGGAATCTGTCGCATTGTTCCT
CCTCCTTCCTGGAAACCGCCGTGCCCACTGAAAGAAAAGCAAGTATGGGAGGGTTCTAAATTTACA
ACACGTGTCCAAAGAGTTGATAAACTTCAGAATCGGAGTTCAATGAAGAAGATTTCAAAGCTTCCT
AATCAAATGAGAAAGAAGAAAAGAAAGTGCATGAAAATGGGAATGGATTCTGTCACTAATGGTATG
GGCGATCCCTGTTCTGCAAGCACTGGAATGAATGAACTTGAGACGTTTGGGTTTGAACCTGGTCCA
GGGTTCACACTAAAAGACTTCCAAAAATATGCTGATGAGTTCAAGGCTCAGTACTTTAAAAAGAGT

FIGURA 10 CONT.

GAAACTTCTACAGACGATAAATGTAAAGTTGATAATTCAATAGACTGCTGGGAGCCAGCACTTGAG
GATGTTGAAGGTGAATATTGGCGGATAGTAGATAAAGCAACTGAAGAGATAGAGGTGCTATATGGT
GCTGACCTTGAAACTGGGGTATTTGGTAGCGGATTCCCCAAGATATCATCAAGTCACAATGCTTCT
TCTTCAGAAGATAAATATGCAAAATCAGGCTGGAACTTAAATAACTTTCCAAGGCTTCCGGGATCC
CTCCTTAAGTACGAGGGCAGTGATATTTCCGGTGTCCTTGTGCCGTGGCTGTATATTGGGATGTGC
TTTTCTTCTTTCTGCTGGCATGTTGAAGATCACCACTTGTATTCGTTGAATTATATGCACTGGGGT
GCACCGAAATTGTGGTATGGTGTTGGGGGGAAGGATGCTGTTAAACTCGAGGAGGCAATGAGGAAG
CATTTGCCTGACCTTTTTGAAGAACAGCCTGACTTACTTCATAAGCTAGTTACACAACTCTCCCCG
TCAAAACTGAAAACCGCAGGAGTACCTGTGCACCGTTGCGTCCAGCATGCTGGAGAGTTTGTCTTG
ACTTTTCCTCGGGCATATCATGCTGGATTTAATTCTGGTTTCAACTGTGCTGAAGCTGTGAATGTA
GCACCCGTTGACTGGTTACCTCATGGGCAGATTGCCATAGAGTTATACTGTCAACAGGGTAGAAAA
ACGTCCATCTCGCATGATAAATTGTTGCTTGGGGCAGCAAGGGAAGTAGTGAAAGCCGACTGGGAG
CTGAACTTACTAAGGAAAAATACTGTAGATAACTTAAGGTGGAAAGCATTCAGTGCGAAGGATGGA
ATTTTGGCAAAAACATTAAAGGCACGCATTGACATGGAACGTACGAGAAGAGAATTTCTGTGCAAC
TCTTCACTTGCATTGAAAATGCACAGTAATTTCGATGCTACCAACGAGAGAGAGTGCTGTATATGC
TTCTTTGATTTGCACCTGTCTGCTGCTGGCTGTCGTTGTTCCCCGGAGAAGTATTCATGTTTGACT
CATGTGAAGGAGTTGTGTTCATGTCCATGGGTTACTAAATATTTTCTATTTCGCTATGATATCGAT
GAACTGAATGTTCTTGTTGAGGCAGTGGAAGGAAAACTGAGTTCTGTATATAGATGGGCGCGTCAA
GATTTAGGATTGGCATTAAGTACAGACGTCTCAGGAAGCAAGATGGAGATAGATGAAGAAGGAAAA
GTCCACAAGGACCCGACTCCACAAACAACTGCACTTTCAGGGAAGGATTTGCAGTTGAAAGTAACA
TCGAAAGAAGTAAGTAAAGAGTTAGAAAAGACTAGTAAATTATCTCATGTTAATTTACTTCTGAAA
GAAAAGGAGGAGCAAATAACGTCAAGTCACTGTATGAAGCCTGTCAAAGAAGAAACTGTTTGTGAT
TCTTCTGATCCAAACGTTTCAGCTTGCCAACCATCTGAAGGAGGCATAATCTGTATGACAGCTGTT
AAGTCCGCAAGTGGTAAAAAGAATTCACAGAGTCTACCCAATGATGTGATACTCCTCAGCGATGAT
GAGTATGACATACCTAGGAAACGAGGTTCTGTGAGAAGAGATGCAATTTCTTCTGGCAAGAAATTA
GAAATACGAGAGAGACCGACCCACGTTTTGGCGTTAGAAGCTTCTGCCAAAATTGCTGCTCCAATT
TGCCAAAGGGAAGGAGATTCGCTGCGTGATACACGAAACACTATATCATTGCCTACTAATGACCAA
AAAACAATGAGAAGGGATGTACCGAGTTCTACTTCGCACGCGGAAGTTAATGCGGAGGCTACTGGG
CTTACTCAAGATATATGTAACAGAATGGCAACTAACAGCCACGGTGGTGGAAAACCTACTAGTTGT
AAATCCAAAAACTCTGGAGGTTTGGCTATAGTGGATGTTGTGGATGGGACAAGAAGTAGTTCAGGT
ACACCGTCTTGTTCGCAAAATAATAGTCCGGATAGGTTCATCCGCCAAAAGGGTCCCCGCATTGCA
AAGGTCGTGAGGAGAATCAATTGCAATGTAGAGCCTTTAAGTTATGGATGTGTGCTGTCTGGAAAA
TCATGGTGCAGCCGGCGAGCAATTTTTCCTAAAGGATTTCGTAGCCGGGTTAAGTACATAAACATT
TTGGATCCAACAAATATGTGCTTCTACATTTCAGAAATTCTGGATGCTGGACGCAACAGTCCATTG
TTCATGGTTTACTTGGAGAGTAATCCCAGTGAGGTGTTCGTTCATATGTCGCCTACAAGATGCTGG
GAGATGGTAAGAGAGAGAGTAAATCAAGAGATAACTAAGCAGCACAAAGCTGGAAAATCAGATCTT
CCTCCTTTGCAACCCTCTGGGAGTCCCGACGGTTTTGAAATGTTTGGATATTCGTCACCCGCAATC
GTACAGGCTATCGAAGCATTAGACGTGAATCGAGTTTGCACAGACTATTGGGATTCCAGGCCTTAT
TCCCGGCCACAAGTTCAGTTCCCTGCAAATCCTCTTCTCAGAGAAGCCAACACAAGTGGCAGATCG
AATGTGGGAAATCTCCAGCTAAACCCTGGACACCATATATCGCCTACTGGAATAAATTCTATTCTT
AAAGTTCTGTTCAAGAAAGCTAGCATGGAGGAACTAAGCTCACTTCAAGAGGTTTTAAGTGAGACT
AACTCAGATATGGTGACCGAACTTGTGAAGGAAGAGATCCAGAACCGTCGCTGA

SEQ ID NO: 96, Proteína **- *Arabidopsis thaliana***

MGTELMRICVKEDSDDLPSVPPGFESYATFTLKRVVPATTSDKAKTPAIESVSATEQAKMEVESDE
AKAARALRRRPWINHSGCDDDGDCAANNDNAASQNPDQNCDVKPALPKGVVRGCEECKDCQKEFED
TLNYIAKIRPEAEKYGICRIVPPPSWKPPCPLKEKQVWEGSKFTTRVQRVDKLQNRSSMKKISKLP
NQMRKKKRKCMKMGMDSVTNGMGDPCSASTGMNELETFGFEPGPGFTLKDFQKYADEFKAQYFKKS

FIGURA 10 CONT.

ETSTDDKCKVDNSIDCWEPALEDVEGEYWRIVDKATEEIEVLYGADLETGVFGSGFPKISSSHNAS
SSEDKYAKSGWNLNNFPRLPGSLLKYEGSDISGVLVPWLYIGMCFSSFCWHVEDHHLYSLNYMHWG
APKLWYGVGGKDAVKLEEAMRKHLPDLFEEQPDLLHKLVTQLSPSKLKTAGVPVHRCVQHAGEFVL
TFPRAYHAGFNSGFNCAEAVNVAPVDWLPHGQIAIELYCQQGRKTSISHDKLLLGAAREVVKADWE
LNLLRKNTVDNLRWKAFSAKDGILAKTLKARIDMERTRREFLCNSSLALKMHSNFDATNERECCIC
FFDLHLSAAGCRCSPEKYSCLTHVKELCSCPWVTKYFLFRYDIDELNVLVEAVEGKLSSVYRWARQ
DLGLALSTDVSGSKMEIDEEGKVHKDPTPQTTALSGKDLQLKVTSKEVSKELEKTSKLSHVNLLLK
EKEEQITSSHCMKPVKEETVCDSSDPNVSACQPSEGGIICMTAVKSASGKKNSQSLPNDVILLSDD
EYDIPRKRGSVRRDAISSGKKLEIRERPTHVLALEASAKIAAPICQREGDSLRDTRNTISLPTNDQ
KTMRRDVPSSTSHAEVNAEATGLTQDICNRMATNSHGGGKPTSCKSKNSGGLAIVDVVDGTRSSSG
TPSCSQNNSPDRFIRQKGPRIAKVVRRINCNVEPLSYGCVLSGKSWCSRRAIFPKGFRSRVKYINI
LDPTNMCFYISEILDAGRNSPLFMVYLESNPSEVFVHMSPTRCWEMVRERVNQEITKQHKAGKSDL
PPLQPSGSPDGFEMFGYSSPAIVQAIEALDVNRVCTDYWDSRPYSRPQVQFPANPLLREANTSGRS
NVGNLQLNPGHHISPTGINSILKVLFKKASMEELSSLQEVLSETNSDMVTELVKEEIQNRR

SEQ ID NO: 97, DNA - *Arabidopsis thaliana*

ATGAATGCTAATGAGCAAACTCGATCCGCCAATGGCATTGGCAATGGCAATGGTGAGTCTATTCCC
GGGATTCCAGATGACTTACGGTGCAAGAGATCGGATGGTAAACAGTGGAGATGCACTGCAATGTCC
ATGGCTGATAAGACTGTTTGTGAGAAGCACTACATCCAAGCAAAGAAGCGGGCGGCTAATTCTGCT
TTCAGGGCGAACCAGAAGAAAGCGAAAAGGCGATCATCGTTAGGCGAAACAGATACGTATTCGGAA
GGGAAGATGGATGATTTCGAGTTACCAGTCACCAGCATTGACCACTATAATAACGGTCTTGCCTCT
GCTTCCAAGAGTAATGGTAGACTAGAGAAGAGACATAATAAAAGCCTGATGCGGTACTCGCCCGAG
ACACCGATGATGAGGAGTTTCTCTCCACGTGTTGCAGTGGATTTGAATGATGACTTGGGTAGAGAT
GTTGTAATGTTTGAAGAGGGCTACAGATCTTATAGGACACCACCATCTGTTGCTGTTATGGATCCG
ACACGAAACAGATCACACCAAAGCACCAGTCCTATGGAATACTCAGCAGCAAGCACAGATGTGTCT
GCAGAGTCTTTGGGGGAAATCTGCCATCAATGCCAGAGAAAAGATAGAGAGAGAATCATTTCTTGC
CTCAAATGCAATCAAAGAGCCTTCTGCCACAATTGTCTATCGGCAAGGTACTCGGAGATATCACTT
GAAGAAGTCGAGAAAGTTTGCCCTGCATGTCGTGGCTTGTGTGATTGCAAATCTTGCCTGCGTTCA
GATAATACAATAAAGGTTCGGATCCGGGAAATACCCGTTTTGGACAAGTTGCAGTATCTTTATCGT
CTATTATCAGCTGTCCTACCAGTCATAAAGCAGATCCATCTTGAACAATGTATGGAAGTTGAACTA
GAGAAGAGGCTTCGTGAAGTTGAGATTGATCTTGTCAGGGCAAGATTGAAAGCAGATGAGCAGATG
TGCTGCAACGTGTGTCGGATACCAGTTGTTGACTACTACCGTCACTGTCCGAACTGCTCATATGAC
CTTTGCCTGAGATGCTGTCAAGATCTACGGGAAGAGTCTTCAGTGACGATTAGTGGGACTAACCAA
AACGTACAAGATAGAAAAGGAGCTCCCAAACTAAAACTAAACTTTTCATACAAGTTTCCTGAGTGG
GAAGCCAACGGTGATGGGAGCATCCCTTGCCCTCCTAAGGAGTATGGAGGCTGCGGTTCACATTCT
TTGAATCTTGCCCGCATTTTCAAGATGAATTGGGTTGCAAAGCTTGTGAAAAATGCTGAGGAGATT
GTTAGTGGCTGCAAATTATCTGATCTTCTGAACCCTGATATGTGTGATTCAAGATTCTGCAAATTT
GCTGAGAGAGAAGAGAGCGGTGACAACTACGTGTACAGCCCGTCGCTTGAAACGATTAAAACTGAT
GGAGTAGCTAAGTTTGAGCAACAATGGGCAGAGGGTCGGCTTGTTACTGTGAAAATGGTACTTGAT
GACTCATCTTGCTCTAGATGGGATCCTGAGACTATTTGGAGGGATATAGACGAGCTTTCGGACGAG
AAACTGAGAGAACATGATCCATTCTTGAAGGCCATTAATTGCTTGGATGGTTTAGAGGTTGATGTA
AGACTTGGGGAGTTTACAAGAGCATATAAAGATGGAAAGAACCAAGAGACAGGTCTTCCGCTATTG
TGGAAGTTAAAGGACTGGCCGAGCCCAAGTGCTTCCGAGGAGTTCATTTTCTACCAAAGACCTGAG
TTTATCAGAAGTTTTCCGTTTCTCGAGTACATTCATCCCCGGTTAGGCCTTCTGAATGTTGCAGCC
AAGTTACCTCATTACTCGCTCCAAAACGATTCAGGTCCAAAGATTTATGTGTCTTGTGGGACGTAC
CAAGAAATCAGTGCTGGCGATTCATTGACTGGTATTCACTACAACATGCGTGACATGGTATACCTA
TTGGTGCACACGTCTGAAGAAACAACATTCGAAAGGGTGAGAAAAACAAAACCTGTTCCAGAGGAA
CCTGACCAGAAGATGAGCGAAAATGAGTCACTTCTTAGCCCTGAGCAGAAATTAAGGGACGGAGAG

FIGURA 10 CONT.

TTACATGATCTATCACTTGGTGAAGCCAGTATGGAGAAGAATGAACCTGAGTTGGCGTTGACTGTG
AATCCAGAGAACTTAACGGAAAACGGTGACAACATGGAATCTTCTTGCACATCTTCATGTGCAGGA
GGAGCCCAGTGGGATGTCTTTCGACGCCAAGACGTCCCAAAGTTGTCCGGGTATTTGCAGAGAACA
TTCCAGAAGCCTGATAATATCCAGACTGATTTTGTGTCACGCCCGTTGTATGAAGGATTGTTCTTA
AATGAACACCACAAGAGACAACTAAGAGACGAGTTTGGAGTTGAGCCATGGACATTTGAGCAACAT
CGTGGTGAGGCTATCTTCATTCCGGCTGGATGTCCGTTCCAAATCACTAATCTTCAGTCGAATATT
CAGGTGGCACTTGACTTCTTGTGCCCTGAAAGCGTTGGAGAGTCAGCAAGACTAGCTGAAGAAATC
CGGTGTTTACCAAACGACCACGAGGCAAAACTTCAGATTCTAGAGATTGGAAAGATATCATTATAC
GCAGCTAGCTCAGCCATTAAAGAGGTTCAGAAACTGGTCTTGGATCCAAAGTTTGGAGCAGAGCTT
GGATTTGAAGACTCTAACTTAACCAAAGCAGTCTCTCACAACTTAGACGAGGCAACCAAGCGGCCG
CAGCAAAACAGCTGCACTTAA

SEQ ID NO: 98, Proteína - *Arabidopsis thaliana*

MNANEQTRSANGIGNGNGESIPGIPDDLRCKRSDGKQWRCTAMSMADKTVCEKHYIQAKKRAANSA
FRANQKKAKRRSSLGETDTYSEGKMDDFELPVTSIDHYNNGLASASKSNGRLEKRHNKSLMRYSPE
TPMMRSFSPRVAVDLNDDLGRDVVMFEEGYRSYRTPPSVAVMDPTRNRSHQSTSPMEYSAASTDVS
AESLGEICHQCQRKDRERIISCLKCNQRAFCHNCLSARYSEISLEEVEKVCPACRGLCDCKSCLRS
DNTIKVRIREIPVLDKLQYLYRLLSAVLPVIKQIHLEQCMEVELEKRLREVEIDLVRARLKADEQM
CCNVCRIPVVDYYRHCPNCSYDLCLRCCQDLREESSVTISGTNQNVQDRKGAPKLKLNFSYKFPEW
EANGDGSIPCPPKEYGGCGSHSLNLARIFKMNWVAKLVKNAEEIVSGCKLSDLLNPDMCDSRFCKF
AEREESGDNYVYSPSLETIKTDGVAKFEQQWAEGRLVTVKMVLDDSSCSRWDPETIWRDIDELSDE
KLREHDPFLKAINCLDGLEVDVRLGEFTRAYKDGKNQETGLPLLWKLKDWPSPSASEEFIFYQRPE
FIRSFPFLEYIHPRLGLLNVAAKLPHYSLQNDSGPKIYVSCGTYQEISAGDSLTGIHYNMRDMVYL
LVHTSEETTFERVRKTKPVPEEPDQKMSENESLLSPEQKLRDGELHDLSLGEASMEKNEPELALTV
NPENLTENGDNMESSCTSSCAGGAQWDVFRRQDVPKLSGYLQRTFQKPDNIQTDFVSRPLYEGLFL
NEHHKRQLRDEFGVEPWTFEQHRGEAIFIPAGCPFQITNLQSNIQVALDFLCPESVGESARLAEEI
RCLPNDHEAKLQILEIGKISLYAASSAIKEVQKLVLDPKFGAELGFEDSNLTKAVSHNLDEATKRP
QQNSCT

SEQ ID NO: 99, DNA - *Arabidopsis thaliana*

ATGAATGCTAATGAGCAAACTCGATCCGCCAATGGCATTGGCAATGGCAATGGTGAGTCTATTCCC
GGGATTCCAGATGACTTACGGTGCAAGAGATCGGATGGTAAACAGTGGAGATGCACTGCAATGTCC
ATGGCTGATAAGACTGTTTGTGAGAAGCACTACATCCAAGCAAAGAAGCGGGCGGCTAATTCTGCT
TTCAGGGCGAACCAGAAGAAAGCGAAAAGGCGATCATCGTTAGGCGAAACAGATACGTATTCGGAA
GGGAAGATGGATGATTTCGAGTTACCAGTCACCAGCATTGACCACTATAATAACGGTCTTGCCTCT
GCTTCCAAGAGTAATGGTAGACTAGAGAAGAGACATAATAAAAGCCTGATGCGGTACTCGCCCGAG
ACACCGATGATGAGGAGTTTCTCTCCACGTGTTGCAGTGGATTTGAATGATGACTTGGGTAGAGAT
GTTGTAATGTTTGAAGAGGGCTACAGATCTTATAGGACACCACCATCTGTTGCTGTTATGGATCCG
ACACGAAACAGATCACACCAAAGCACCAGTCCTATGGAATACTCAGCAGCAAGCACAGATGTGTCT
GCAGAGTCTTTGGGGGAAATCTGCCATCAATGCCAGAGAAAAGATAGAGAGAGAATCATTTCTTGC
CTCAAATGCAATCAAAGAGCCTTCTGCCACAATTGTCTATCGGCAAGGTACTCGGAGATATCACTT
GAAGAAGTCGAGAAAGTTTGCCCTGCATGTCGTGGCTTGTGTGATTGCAAATCTTGCCTGCGTTCA
GATAATACAATAAAGGTTCGGATCCGGGAAATACCCGTTTTGGACAAGTTGCAGTATCTTTATCGT
CTATTATCAGCTGTCCTACCAGTCATAAAGCAGATCCATCTTGAACAATGTATGGAAGTTGAACTA
GAGAAGAGGCTTCGTGAAGTTGAGATTGATCTTGTCAGGGCAAGATTGAAAGCAGATGAGCAGATG
TGCTGCAACGTGTGTCGGATACCAGTTGTTGACTACTACCGTCACTGTCCGAACTGCTCATATGAC
CTTTGCCTGAGATGCTGTCAAGATCTACGGGAAGAGTCTTCAGTGACGATTAGTGGGACTAACCAA
AACGTACAAGATAGAAAAGGAGCTCCCAAACTAAAACTAAACTTTTCATACAAGTTTCCTGAGTGG

FIGURA 10 CONT.

GAAGCCAACGGTGATGGGAGCATCCCTTGCCCTCCTAAGGAGTATGGAGGCTGCGGTTCACATTCT
TTGAATCTTGCCCGCATTTTCAAGATGAATTGGGTTGCAAAGCTTGTGAAAAATGCTGAGGAGATT
GTTAGTGGCTGCAAATTATCTGATCTTCTGAACCCTGATATGTGTGATTCAAGATTCTGCAAATTT
GCTGAGAGAGAAGAGAGCGGTGACAACTACGTGTACAGCCCGTCGCTTGAAACGATTAAAACTGAT
GGAGTAGCTAAGTTTGAGCAACAATGGGCAGAGGGTCGGCTTGTTACTGTGAAAATGGTACTTGAT
GACTCATCTTGCTCTAGATGGGATCCTGAGACTATTTGGAGGGATATAGACGAGCTTTCGGACGAG
AAACTGAGAGAACATGATCCATTCTTGAAGGCCATTAATTGCTTGGATGGTTTAGAGGTTGATGTA
AGACTTGGGGAGTTTACAAGAGCATATAAAGATGGAAAGAACCAAGAGACAGGTCTTCCGCTATTG
TGGAAGTTAAAGGACTGGCCGAGCCCAAGTGCTTCCGAGGAGTTCATTTTCTACCAAAGACCTGAG
TTTATCAGAAGTTTTCCGTTTCTCGAGTACATTCATCCCCGGTTAGGCCTTCTGAATGTTGCAGCC
AAGTTACCTCATTACTCGCTCCAAAACGATTCAGGTCCAAAGATTTATGTGTCTTGTGGGACGTAC
CAAGAAATCAGTGCTGGCGATTCATTGACTGGTATTCACTACAACATGCGTGACATGGTATACCTA
TTGGTGCACACGTCTGAAGAAACAACATTCGAAAGGGTGAGAAAAACAAAACCTGTTCCAGAGGAA
CCTGACCAGAAGATGAGCGAAAATGAGTCACTTCTTAGCCCTGAGCAGAAATTAAGGGACGGAGAG
TTACATGATCTATCACTTGGTGAAGCCAGTATGGAGAAGAATGAACCTGAGTTGGCGTTGACTGTG
AATCCAGAGAACTTAACGGAAAACGGTGACAACATGGAATCTTCTTGCACATCTTCATGTGCAGGA
GGAGCCCAGTGGGATGTCTTTCGACGCCAAGACGTCCCAAAGTTGTCCGGGTATTTGCAGAGAACA
TTCCAGAAGCCTGATAATATCCAGACTGATTTTGTGTCACGCCCGTTGTATGAAGGATTGTTCTTA
AATGAACACCACAAGAGACAACTAAGAGACGAGTTTGGAGTTGAGCCATGGACATTTGAGCAACAT
CGTGGTGAGGCTATCTTCATTCCGGCTGGATGTCCGTTCCAAATCACTAATCTTCAGTCGAATATT
CAGGTGGCACTTGACTTCTTGTGCCCTGAAAGCGTTGGAGAGTCAGCAAGACTAGCTGAAGAAATC
CGGTGTTTACCAAACGACCACGAGGCAAAACTTCAGATTCTAGAGATTGGAAAGATATCATTATAC
GCAGCTAGCTCAGCCATTAAAGAGGTTCAGAAACTGGTCTTGGATCCAAAGTTTGGAGCAGAGCTT
GGATTTGAAGACTCTAACTTAACCAAAGCAGTCTCTCACAACTTAGACGAGGCAACCAAGCGGCCG
CAGCAAAACAGCTGCACTTAA

SEQ ID NO: 100, Proteína - *Arabidopsis thaliana*

MNANEQTRSANGIGNGNGESIPGIPDDLRCKRSDGKQWRCTAMSMADKTVCEKHYIQAKKRAANSA
FRANQKKAKRRSSLGETDTYSEGKMDDFELPVTSIDHYNNGLASASKSNGRLEKRHNKSLMRYSPE
TPMMRSFSPRVAVDLNDDLGRDVVMFEEGYRSYRTPPSVAVMDPTRNRSHQSTSPMEYSAASTDVS
AESLGEICHQCQRKDRERIISCLKCNQRAFCHNCLSARYSEISLEEVEKVCPACRGLCDCKSCLRS
DNTIKVRIREIPVLDKLQYLYRLLSAVLPVIKQIHLEQCMEVELEKRLREVEIDLVRARLKADEQM
CCNVCRIPVVDYYRHCPNCSYDLCLRCCQDLREESSVTISGTNQNVQDRKGAPKLKLNFSYKFPEW
EANGDGSIPCPPKEYGGCGSHSLNLARIFKMNWVAKLVKNAEEIVSGCKLSDLLNPDMCDSRFCKF
AEREESGDNYVYSPSLETIKTDGVAKFEQQWAEGRLVTVKMVLDDSSCSRWDPETIWRDIDELSDE
KLREHDPFLKAINCLDGLEVDVRLGEFTRAYKDGKNQETGLPLLWKLKDWPSPSASEEFIFYQRPE
FIRSFPFLEYIHPRLGLLNVAAKLPHYSLQNDSGPKIYVSCGTYQEISAGDSLTGIHYNMRDMVYL
LVHTSEETTFERVRKTKPVPEEPDQKMSENESLLSPEQKLRDGELHDLSLGEASMEKNEPELALTV
NPENLTENGDNMESSCTSSCAGGAQWDVFRRQDVPKLSGYLQRTFQKPDNIQTDFVSRPLYEGLFL
NEHHKRQLRDEFGVEPWTFEQHRGEAIFIPAGCPFQITNLQSNIQVALDFLCPESVGESARLAEEI
RCLPNDHEAKLQILEIGKISLYAASSAIKEVQKLVLDPKFGAELGFEDSNLTKAVSHNLDEATKRP
QQNSCT

SEQ ID NO: 101, DNA - *Arabidopsis thaliana*

ATGCAGGTCAATTTTGATGAAACTTGTGATTCAGTGATTAGAATGAATGCTAATGAGCAAACTCGA
TCCGCCAATGGCATTGGCAATGGCAATGGTGAGTCTATTCCCGGGATTCCAGATGACTTACGGTGC
AAGAGATCGGATGGTAAACAGTGGAGATGCACTGCAATGTCCATGGCTGATAAGACTGTTTGTGAG
AAGCACTACATCCAAGCAAAGAAGCGGGCGGCTAATTCTGCTTTCAGGGCGAACCAGAAGAAAGCG

FIGURA 10 CONT.

AAAAGGCGATCATCGTTAGGCGAAACAGATACGTATTCGGAAGGGAAGATGGATGATTTCGAGTTA
CCAGTCACCAGCATTGACCACTATAATAACGGTCTTGCCTCTGCTTCCAAGAGTAATGGTAGACTA
GAGAAGAGACATAATAAAAGCCTGATGCGGTACTCGCCCGAGACACCGATGATGAGGAGTTTCTCT
CCACGTGTTGCAGTGGATTTGAATGATGACTTGGGTAGAGATGTTGTAATGTTTGAAGAGGGCTAC
AGATCTTATAGGACACCACCATCTGTTGCTGTTATGGATCCGACACGAAACAGATCACACCAAAGC
ACCAGTCCTATGGAATACTCAGCAGCAAGCACAGATGTGTCTGCAGAGTCTTTGGGGGAAATCTGC
CATCAATGCCAGAGAAAAGATAGAGAGAGAATCATTTCTTGCCTCAAATGCAATCAAAGAGCCTTC
TGCCACAATTGTCTATCGGCAAGGTACTCGGAGATATCACTTGAAGAAGTCGAGAAAGTTTGCCCT
GCATGTCGTGGCTTGTGTGATTGCAAATCTTGCCTGCGTTCAGATAATACAATAAAGGTTCGGATC
CGGGAAATACCCGTTTTGGACAAGTTGCAGTATCTTTATCGTCTATTATCAGCTGTCCTACCAGTC
ATAAAGCAGATCCATCTTGAACAATGTATGGAAGTTGAACTAGAGAAGAGGCTTCGTGAAGTTGAG
ATTGATCTTGTCAGGGCAAGATTGAAAGCAGATGAGCAGATGTGCTGCAACGTGTGTCGGATACCA
GTTGTTGACTACTACCGTCACTGTCCGAACTGCTCATATGACCTTTGCCTGAGATGCTGTCAAGAT
CTACGGGAAGAGTCTTCAGTGACGATTAGTGGGACTAACCAAAACGTACAAGATAGAAAAGGAGCT
CCCAAACTAAAACTAAACTTTTCATACAAGTTTCCTGAGTGGGAAGCCAACGGTGATGGGAGCATC
CCTTGCCCTCCTAAGGAGTATGGAGGCTGCGGTTCACATTCTTTGAATCTTGCCCGCATTTTCAAG
ATGAATTGGGTTGCAAAGCTTGTGAAAAATGCTGAGGAGATTGTTAGTGGCTGCAAATTATCTGAT
CTTCTGAACCCTGATATGTGTGATTCAAGATTCTGCAAATTTGCTGAGAGAGAAGAGAGCGGTGAC
AACTACGTGTACAGCCCGTCGCTTGAAACGATTAAAACTGATGGAGTAGCTAAGTTTGAGCAACAA
TGGGCAGAGGGTCGGCTTGTTACTGTGAAAATGGTACTTGATGACTCATCTTGCTCTAGATGGGAT
CCTGAGACTATTTGGAGGGATATAGACGAGCTTTCGGACGAGAAACTGAGAGAACATGATCCATTC
TTGAAGGCCATTAATTGCTTGGATGGTTTAGAGGTTGATGTAAGACTTGGGGAGTTTACAAGAGCA
TATAAAGATGGAAAGAACCAAGAGACAGGTCTTCCGCTATTGTGGAAGTTAAAGGACTGGCCGAGC
CCAAGTGCTTCCGAGGAGTTCATTTTCTACCAAAGACCTGAGTTTATCAGAAGTTTTCCGTTTCTC
GAGTACATTCATCCCCGGTTAGGCCTTCTGAATGTTGCAGCCAAGTTACCTCATTACTCGCTCCAA
AACGATTCAGGTCCAAAGATTTATGTGTCTTGTGGGACGTACCAAGAAATCAGTGCTGGCGATTCA
TTGACTGGTATTCACTACAACATGCGTGACATGGTATACCTATTGGTGCACACGTCTGAAGAAACA
ACATTCGAAAGGGTGAGAAAAACAAAACCTGTTCCAGAGGAACCTGACCAGAAGATGAGCGAAAAT
GAGTCACTTCTTAGCCCTGAGCAGAAATTAAGGGACGGAGAGTTACATGATCTATCACTTGGTGAA
GCCAGTATGGAGAAGAATGAACCTGAGTTGGCGTTGACTGTGAATCCAGAGAACTTAACGGAAAAC
GGTGACAACATGGAATCTTCTTGCACATCTTCATGTGCAGGAGGAGCCCAGTGGGATGTCTTTCGA
CGCCAAGACGTCCCAAAGTTGTCCGGGTATTTGCAGAGAACATTCCAGAAGCCTGATAATATCCAG
ACTGATTTTGTGTCACGCCCGTTGTATGAAGGATTGTTCTTAAATGAACACCACAAGAGACAACTA
AGAGACGAGTTTGGAGTTGAGCCATGGACATTTGAGCAACATCGTGGTGAGGCTATCTTCATTCCG
GCTGGATGTCCGTTCCAAATCACTAATCTTCAGTCGAATATTCAGGTGGCACTTGACTTCTTGTGC
CCTGAAAGCGTTGGAGAGTCAGCAAGACTAGCTGAAGAAATCCGGTGTTTACCAAACGACCACGAG
GCAAAACTTCAGATTCTAGAGATTGGAAAGATATCATTATACGCAGCTAGCTCAGCCATTAAAGAG
GTTCAGAAACTGGTCTTGGATCCAAAGTTTGGAGCAGAGCTTGGATTTGAAGACTCTAACTTAACC
AAAGCAGTCTCTCACAACTTAGACGAGGCAACCAAGCGGCCGCAGCAAAACAGCTGCACTTAA

SEQ ID NO: 102, Proteína - *Arabidopsis thaliana*

MQVNFDETCDSVIRMNANEQTRSANGIGNGNGESIPGIPDDLRCKRSDGKQWRCTAMSMADKTVCE
KHYIQAKKRAANSAFRANQKKAKRRSSLGETDTYSEGKMDDFELPVTSIDHYNNGLASASKSNGRL
EKRHNKSLMRYSPETPMMRSFSPRVAVDLNDDLGRDVVMFEEGYRSYRTPPSVAVMDPTRNRSHQS
TSPMEYSAASTDVSAESLGEICHQCQRKDRERIISCLKCNQRAFCHNCLSARYSEISLEEVEKVCP
ACRGLCDCKSCLRSDNTIKVRIREIPVLDKLQYLYRLLSAVLPVIKQIHLEQCMEVELEKRLREVE
IDLVRARLKADEQMCCNVCRIPVVDYYRHCPNCSYDLCLRCCQDLREESSVTISGTNQNVQDRKGA
PKLKLNFSYKFPEWEANGDGSIPCPPKEYGGCGSHSLNLARIFKMNWVAKLVKNAEEIVSGCKLSD

FIGURA 10 CONT.

LLNPDMCDSRFCKFAEREESGDNYVYSPSLETIKTDGVAKFEQQWAEGRLVTVKMVLDDSSCSRWD
PETIWRDIDELSDEKLREHDPFLKAINCLDGLEVDVRLGEFTRAYKDGKNQETGLPLLWKLKDWPS
PSASEEFIFYQRPEFIRSFPFLEYIHPRLGLLNVAAKLPHYSLQNDSGPKIYVSCGTYQEISAGDS
LTGIHYNMRDMVYLLVHTSEETTFERVRKTKPVPEEPDQKMSENESLLSPEQKLRDGELHDLSLGE
ASMEKNEPELALTVNPENLTENGDNMESSCTSSCAGGAQWDVFRRQDVPKLSGYLQRTFQKPDNIQ
TDFVSRPLYEGLFLNEHHKRQLRDEFGVEPWTFEQHRGEAIFIPAGCPFQITNLQSNIQVALDFLC
PESVGESARLAEEIRCLPNDHEAKLQILEIGKISLYAASSAIKEVQKLVLDPKFGAELGFEDSNLT
KAVSHNLDEATKRPQQNSCT

SEQ ID NO: 103, DNA - *Arabidopsis thaliana*

ATGGAGGGTGAAGTTGCTACTAACGGAGTGATTCTCAAGCATAATGGTGTAAAGGATATTTCTTTA
GAAACTTGTTGGCCAGAGAAGAAGAAACCGGTTGAGGCAACGAGTCTTAGCAGTGGTTCATCAGAC
ATAGAAGAAGAGATTAGTGTGGAATGTCCTAAGCGAGTGGCTAATCAAAGACGGAAACGGTCTAAA
GCTGATGAGATTAAAACGAAATCTTCAAGAAAGAGGAAATGTGATGATGAGAACAAGTGTGAAGAG
AATGAGAAGAAGCAGAGGAGCTCTGTTAAAAAGAGAGCAACTACATGGAAAGAAGAAGAAGTTGTT
GTTGATGATGAAAAGAAATGCGAACAACAACTCCAACTAGTTCCTTCCAGCAAGGCAACATCTCGA
AGCAGGAGCAAGAAATCAGTCAGTGTGGATACTTGGTTGGTTAACAACGAGATTGATGTTAGCGCT
TTGAGTTCGCGTTCTGAGTCTGAACTTTCGGATTCTTATCTGAAAACCGAGTACTTTAATGACTGC
AGAAGCATGACAAGGAGTTTAAAGGCAAATTTGGGAGAGCTTGCAATTTGCCATCAATGCTCTAAA
GGGGAAAGAAGATATCTTTTCATTTGCACCTTTTGTGAAGTGAGGTTGTATTGTTTTCCATGCATA
AAGAAATGGTATCCCCATTTGTCCACGGATGATATCCTTGAAAAATGTCCCTTTTGCCGAGGAACT
TGCAACTGCTGCACATGTTTGCACTCTAGTGGTTTAATCGAGACATCTAAGAGAAAGCTCGATAAA
TACGAAAGATTTTATCATCTTCGCTTCTTGATTGTGGCGATGCTTCCTTTCCTGAAAAAGTTATGC
AAAGCACAAGATCAAGAAATCGAAACCGAGGCTAAGGTTCAAGATTCAATGGCTTCTCAAGTTGAT
ATATCCGAGAGTCTATGCTCAAATGAAGAGCGTGTCTTCTGTAACCACTGTGCAACATCAATTGTT
GACTTGCATCGAAGCTGTCCAAAGTGTTCATATGAATTGTGTCTGAACTGTTGCCAAGAGATCCGT
GGAGGTTGGCTTTCTGACCGCCCGGAATGTCAGCTACAATTCGAGTACAGAGGGACCCGGTATATA
CATGGCGAAGCTGCAGAACCGAGCTCTTCTTCTGTTTCCGAGGATGAAACTAAAACTCCATCCATC
AAGTGGAATGCTGATGAAAATGGAAGCATACGGTGTGCGCCAAAAGAGCTAGGAGGTTGCGGGGAC
TCTGTGCTGGAGCTTAAGCGAATCTTACCAGTGACTTGGATGTCAGATTTGGAGCAAAAGGCAGAA
ACATTCTTAGCCTCTTATAGTATCAAACCGCCAATGTCATATTGTAGATGTTCTTCTGATATGAGT
AGTATGAAGAGGAAGGCAGCTTCGAGGGACGGATCTAGTGACAATTACCTTTATTCTCCTGATTCT
TTGGATGTCCTGAAGCAAGAAGAACTTCTGCATTTCCAAGAGCATTGGTCCAAAGGTGAACCAGTG
ATTGTTCGAAATGCTCTTAACAACACAGCTGGTTTAAGCTGGGAGCCAATGGTGATGTGGCGGGCT
TTATGTGAGAATGTTGATTCAGCAATTAGCTCTAATATGTCTGATGTCAAAGCTATCGATTGTTTA
GCTAATTGTGAGGTGAAGATAAATACTCTATGTTTCTTTGAAGGATATAGCAAAGGAAGAACATAT
GAAAACTTCTGGCCTGAGATGCTGAAGCTGAAGGATTGGCCTCCTTCAGACAAGTTTGAAAACCTT
TTACCTCGTCACTGCGATGAGTTCATTTCCGCATTACCATTTCAAGAATACAGCGATCCTAGATCA
GGAATTCTCAACATTGCTACAAAACTTCCTGAAGGACTTCTTAAACCAGATCTTGGTCCAAAAACT
TATGTTGCATATGGGACTTCAGATGAGCTCGGTAGAGGAGATTCAGTCACTAAGCTTCACTGTGAT
ATGTCAGATGCTGTAAGTTTCAGAATTGTTAATATTCTGATGCATACTGCTGAAGTAACACTAAGT
GAGGAGCAAAGGTCTGCAATTGCAGATTTGAAACAGAAACATAAGCAACAGAATGAGAAAGAGCTT
CAGGAGCAAAATGGTTTAGAGGAAGAAGAGGTTGTGAGTGATGAGATTGTGGTTTATGATGAAACA
AGTGGTGCACTTTGGGACATCTTTAAAAGAGAAGATGTTCCTAAGCTTGAAGAGTATCTAAGGAAG
CATTGCATAGAATTCAGACATACTTATTGCTCTCGTGTTACAAAGGTGTATCATCCAATACATGAT
CAGTCATATTTCTTAACCGTGGAGCATAAAAGAAAACTCAAGGCAGAATTTGGGATTGAGCCATGG
ACATTTGTGCAAAAGCTAGGAGAAGCAGTGTTTATTCCAGCAGGTTGTCCTCATCAAGTTCGGAAT

FIGURA 10 CONT.

CTCAAGTCGTGCACAAAGGTAGCTGTTGATTTTGTTTCCCCTGAAAACATCGATGAGTGTCTCCGT
TTGACTGATGAGTTTCGACAACTCCCTAAGAACCATAAAGCCAGAGAAGACAAACTTGAGATAAAA
AAGATGGTGATCTATGCAGTTGAACAAGCTCTGAAAGAAGTGGAGACATTATTGCTAGATCGCAGC
TAA

SEQ ID NO: 104, Proteína - *Arabidopsis thaliana*

MEGEVATNGVILKHNGVKDISLETCWPEKKKPVEATSLSSGSSDIEEEISVECPKRVANQRRKRSK
ADEIKTKSSRKRKCDDENKCEENEKKQRSSVKKRATTWKEEEVVVDDEKKCEQQLQLVPSSKATSR
SRSKKSVSVDTWLVNNEIDVSALSSRSESELSDSYLKTEYFNDCRSMTRSLKANLGELAICHQCSK
GERRYLFICTFCEVRLYCFPCIKKWYPHLSTDDILEKCPFCRGTCNCCTCLHSSGLIETSKRKLDK
YERFYHLRFLIVAMLPFLKKLCKAQDQEIETEAKVQDSMASQVDISESLCSNEERVFCNHCATSIV
DLHRSCPKCSYELCLNCCQEIRGGWLSDRPECQLQFEYRGTRYIHGEAAEPSSSSVSEDETKTPSI
KWNADENGSIRCAPKELGGCGDSVLELKRILPVTWMSDLEQKAETFLASYSIKPPMSYCRCSSDMS
SMKRKAASRDGSSDNYLYSPDSLDVLKQEELLHFQEHWSKGEPVIVRNALNNTAGLSWEPMVMWRA
LCENVDSAISSNMSDVKAIDCLANCEVKINTLCFFEGYSKGRTYENFWPEMLKLKDWPPSDKFENL
LPRHCDEFISALPFQEYSDPRSGILNIATKLPEGLLKPDLGPKTYVAYGTSDELGRGDSVTKLHCD
MSDAVSFRIVNILMHTAEVTLSEEQRSAIADLKQKHKQQNEKELQEQNGLEEEEVVSDEIVVYDET
SGALWDIFKREDVPKLEEYLRKHCIEFRHTYCSRVTKVYHPIHDQSYFLTVEHKRKLKAEFGIEPW
TFVQKLGEAVFIPAGCPHQVRNLKSCTKVAVDFVSPENIDECLRLTDEFRQLPKNHKAREDKLEIK
KMVIYAVEQALKEVETLLLDRS

SEQ ID NO: 105, DNA - *Arabidopsis thaliana*

ATGGAAAATCCTCCATTAGAATCTGAGATCAAAGAGGATATGTCTTTGAAGAATCATCCTCCAGAC
AAGGATAAGGACAAAGACACTATCATGGAACAACCTAGTAGTCCACGTCATCGGAAGGTTGTTGCT
AGATGGTTACCTGATGAAGCACAGAGGCCAATTATCAATGACGCTCCCGTATTCACTCCATCACTA
GAGGAGTTTGTAGATCCACTTGCGTATATCGAGAAGATACGTCCACTAGCAGAGCCGTATGGTATC
TGTCGAATCATCCCACCATCGACATGGAAGCCTCCTTGCCGGCTTAAGGAGAAGAGTATATGGGAG
CAGACAAAGTTTCCTACTCGGATTCAAACCGTAGACCTGCTTCAGAACCGTGAGCCGATGAAGAAG
AAACCGAAGAGTAGAAAGCGGAAACGGAGAAGAAACTCGAGAATGGGTTCGTCTAAGAGACGATCT
GGTTCTTCTCCCGCTGAATCGACTTCATCACCCGAGGCTGAGGAGAAGTTTGGTTTCAATTCTGGT
TCAGACTTCACTCTTGATGAGTTTGAGAAATATGCTCTGCATTTTAAAGACTCGTACTTTGAAAAG
AAAGACTCTGGTGGAGATATAGTAAAATGGACACCGTCTGTGGATGATATAGAAGGGGAATACTGG
CGGATAGTTGAGCAACCAACAGATGAAGTGGAGGTTTACTATGGAGCTGACTTGGAGAATGGGGTG
CTTGGAAGCGGGTTTTATAAAAGAGCCGAGAAGTTTACCGGTAGCGATATGGAGCAGTACACATTA
TCTGGCTGGAACTTGAATAACTTACCACGGCTTCCTGGCTCTGTACTCTCTTTTGAAGATTGTGAC
ATATCCGGAGTTCTAGTTCCATGGCTATATGTGGGAATGTGTTTTTCATCATTTTGTTGGCATGTG
GAGGACCATCATTTATATTCACTTAACTATCATCACTTTGGGGAGCCAAAAGTATGGTATGGTGTT
CCAGGAAGCAATGCAACCGCTCTCGAGAAGGCGATGAGGAAACATTTACCTGACTTGTTCGAAGAC
AGCCTGATCTACTACATGGCCTGGGCGTATCATGCTGGATTCAACTGCGGTTTCAACTGCGCAGAA
GCGGTTAATGTGGCTCCGGTTGATTGGTTGGCTCATGGACAGAACGCTGTGGAACTATATAGTAAA
GAGACAAGGAAGACTTCTCTGTCTCACGACAAGCTTCTTCTTGGAGCAGCTTATGAAGCGGTGAAG
GCTCTTTGGGAACTCTCAGCTTCTGAGGGAAAGGAAAATACAACAAATTTGAGATGGAAGAGTTTC
TGTGGGAAGAACGGGACACTTACCAACGCGATACAAGCTCGGTTACAAATGGAAGAGGGGAGAATT
ACAGCTCTTGGTAGGGATTCTTCAAGCTTGAAGAAGATGGAGAAGGACTTTGATTCAAACTGTGAA
AGGGAATGTTTCTCATGCTTTTATGATTTGCATCTCTCGGCTTCTGGCTGCAAGTGCTCTCCCGAA
GAATACGCGTGCCTTAAACATGCAGATGATCTTTGTTCATGTGATGTGAAAGATGGATTTATTCTT
CTCCGGTACACAATGGATGAGTTAAGCTCGTTGGTCAGAGCATTGGAAGGGGAATCAGATGATTTA
AAGATATGGGCTTCCAAGGTTTTAGGCATTGAACATAGTGATGAAGATCAGACAAAGACTAGTTCA

FIGURA 10 CONT.

GTTATCAGTGAGGAGAAGAAGTTAAAGGAAGGTTCTTTTGATCTGAACATTGACTTAGAGATGGAT
TATCAAGAAGACGTTAAAGAAGAAGCCAGCACTAGTGGCGGCGAGTTAACCGCCTCAGAGAACCTT
GGTGTATCGGTCGAGCCTATAAACCTCGGGTTCTTGATTTTTGGAAAGCTTTGGTGCAATAAGTAT
GCTATATTCCCAAAAGGATTCAGGAGTCGTGTCAAGTTCTACAATGTTCTTGATCCAACAAGAATG
AGCAATTACATCTCTGAGGTTTTGGATGCAGGACTCATGGGTCCATTGTTCAGGGTTACTCTGGAA
GAATCTCCAGACGAGAGCTTCTTCAATGTCTCTGCACAACAATGCTGGGAAATGGTGATGCGGAGA
GTTAAAGATACATCAACAAGTCTTGGTCTTCCTATTTTACCACAGTTTGAGAGTATCAACGGGCTT
CAAATGTTTGGTTTCCTCTCACCCTCTATAGTTCAGGCCATTGAAGCTCTTGACCCAAACCATCGA
CTAGTTGAGTACTGGAACCACAAGAACCAAACTTCATCAGACTCAAAAGATCACTTCATATCATCA
AACTGTTCCGCAAGTTTAACCAAAGGAAAACTTTTCGGAGTAGATTTGATGTAA

SEQ ID NO: 106, Proteína - *Arabidopsis thaliana*

MENPPLESEIKEDMSLKNHPPDKDKDKDTIMEQPSSPRHRKVVARWLPDEAQRPIINDAPVFTPSL
EEFVDPLAYIEKIRPLAEPYGICRIIPPSTWKPPCRLKEKSIWEQTKFPTRIQTVDLLQNREPMKK
KPKSRKRKRRRNSRMGSSKRRSGSSPAESTSSPEAEEKFGFNSGSDFTLDEFEKYALHFKDSYFEK
KDSGGDIVKWTPSVDDIEGEYWRIVEQPTDEVEVYYGADLENGVLGSGFYKRAEKFTGSDMEQYTL
SGWNLNNLPRLPGSVLSFEDCDISGVLVPWLYVGMCFSSFCWHVEDHHLYSLNYHHFGEPKVWYGV
PGSNATALEKAMRKHLPDLFEDSLIYYMAWAYHAGFNCGFNCAEAVNVAPVDWLAHGQNAVELYSK
ETRKTSLSHDKLLLGAAYEAVKALWELSASEGKENTTNLRWKSFCGKNGTLTNAIQARLQMEEGRI
TALGRDSSSLKKMEKDFDSNCERECFSCFYDLHLSASGCKCSPEEYACLKHADDLCSCDVKDGFIL
LRYTMDELSSLVRALEGESDDLKIWASKVLGIEHSDEDQTKTSSVISEEKKLKEGSFDLNIDLEMD
YQEDVKEEASTSGGELTASENLGVSVEPINLGFLIFGKLWCNKYAIFPKGFRSRVKFYNVLDPTRM
SNYISEVLDAGLMGPLFRVTLEESPDESFFNVSAQQCWEMVMRRVKDTSTSLGLPILPQFESINGL
QMFGFLSPSIVQAIEALDPNHRLVEYWNHKNQTSSDSKDHFISSNCSASLTKGKLFGVDLM

SEQ ID NO: 107, DNA - *Arabidopsis thaliana*

ATGGATTCTGGAGTTAAATTGGAGCATATGAATTGTTTCCAATTATCTTATCAATATTCGTGGACG
ACTAGGAAGAAGAGAACTCTGAAACCATTCATGTCAAAAGGTTCTAGCCCTAGTAGTAGTAGTGAT
AGTAGGAAACGGAAGCTTTCGAGAGCTGAAGATAGCGATGATTCAGCGGTTAAACGGAATGCGAAA
CGAAGGAGAAAGATATGCAAAGTAGAAGAATATTATGAAGATGATGATTGCATATTAAGTGATTGG
GTTCAGAGGAATACTGCGAAAAGGATTGATAAGCGCAACGAAGAAGTTGAAGTTATGGTTAAGATA
GAGTCAGGTGATGATTGCACAATAGGTAAGTGGTTTTCTGATGTTAGTAGTAAAAGGAAAGATAAA
CGACAAGTAGAAGTTGATGAAGATGAGGAATGGGAAGAAGAAGTTACACTGTGTTCTAAGATCAAG
GCAACGTCGTCGCGAAGCAGAACCCATAGTTTAAGTGCTAACAGTCCAGAAAATGTTACAGATGTT
ATTAGTCCTTGTCGTTCACGGTCTCCAGCTTCTAATGTATCAGATTCCATTCAGAAAAATGACTGC
ACCAGTAGTAGAAAACAGAGTGGCCCAATTTGTCATCAGTGCTTGAAAGGGGAAAGGATAACACTT
CTCATTTGTAGTGAATGTGAGAAGACAATGTTTTGTCTTCAATGTATAAGGAAATGGTATCCAAAT
CTATCTGAGGATGATGTTGTTGAGAAATGTCCTTTATGCCGTCAAAATTGTAATTGTAGCAAATGC
TTGCATTTGAATGGTTTAATTGAGACATCGAAGAGGGAACTCGCGAAATCTGAAAGACGTCATCAT
CTCCAGTATTTGATTACGTTGATGCTTCCTTTTCTGAATAAGTTATCCATATTCCAGAAGCTAGAG
ATTGAATTTGAGGCAACTGTTCAAGGAAAACTGCCTTCTGAAGTTGAAATAACTGCGGCTATAAGC
TACACTGATGAACGTGTATACTGTGATCATTGTGCAACTTCAATTGTTGACTTGCATCGAAGCTGT
CCAAAGTGCTCTTATGAACTTTGTTTGAAGTGCTGTCAAGAGATTCGTGAAGGATCGCTTTCTGAA
CGCCCTGAAATGAAGTTCCATTATGTTGATAGAGGACATCGATATATGCATGGTTTAGATGCTGCG
GAACCGAGCTTGTCTTCCACTTTTGAAGACGAAGAAGCTAACCCATCTGACGCTAAGTGGAGCCTT
GGTGAAAATGGAAGCATCACTTGTGCACCAGAAAAGCTAGGTGGTTGTGGTGAACGGATGCTAGAG
CTTAGGCGGATCCTACCACTCACGTGGATGTCAGATTTAGAGCACAAAGCAGAGACTTTCTTATCA
TCATACAATATAAGCCCTAGAATGTTGAACTGTAGGTGTTCTTCTTTGGAGACAGAGCTGACAAGG

FIGURA 10 CONT.

AAATCAGCTTCGAGGACAACATCAAGTGACAACTACCTTTTCTGTCCTGAATCTCTCGGTGTCTTG
AAGGAAGAAGAGCTTCTTCATTTTCAAGAGCATTGGGCAAAAGGCGAACCAGTAATCGTTAGGAAC
GCTCTTGACAACACACCTGGTTTAAGCTGGGAGCCGATGGTTATGTGGCGGGCTTTATGCGAAAAT
GTGAATTCAACATCAAGCTCTGAGATGTCCCAAGTCAAAGCAATTGATTGCTTAGCTAATTGTGAG
GTGGAGATCAATACTCGTCAGTTCTTTGAAGGTTATAGCAAAGGAAGAACATATGAAAATTTCTGG
CCTGAGATGTTGAAGCTTAAGGACTGGCCTCCTTCTGATAAGTTTGAAGATCTTCTACCTCGTCAC
TGTGATGAGTTCATATCCGCGTTACCTTTCCAGGAATATAGTGATCCTAGAACTGGAATTCTTAAC
ATTGCAACAAAGCTTCCTGAAGGATTTATCAAACCAGATTTAGGTCCAAAAACTTACATCGCCTAT
GGGATTCCAGATGAGCTTGGTAGAGGCGATTCCGTGACTAAGCTTCACTGCGATATGTCAGACGCG
GTGAATATTCTAACGCATACGGCCGAAGTGACCCTAAGTCAAGAACAAATATCCTCAGTCAAAGCA
CTGAAACAGAAACACAAGTTACAGAATAAGGTCGATAAACAGAGTACTGAAGACTGTAACGAAAAG
GAAGAAGAAGAAGAAGAAGAATTGAACATGCCAGAGATTTCGAGCAATGAAAATGAAGAAACGGGC
AGTGCTCTTTGGGACATATTTAGGAGAGAAGACGTTCCCAAGTTAGAAGAGTATCTAAGAAAGCAT
TGCAAAGAATTTAGACACACTTATTGTAGTCCAGTTACAAAGGTTTATCATCCGATACATGACCAA
TCATGCTATTTAACCTTAGAGCACAAAAGAAAACTCAAGGCGGAATATGGGATCGAGCCATGGACA
TTTGTGCAAAAGCTAGGTGAAGCGGTGTTTATTCCCGCGGGTTGTCCTCATCAAGTTCGGAATCTA
AAGTCGTGCACAAAAGTCGCAGTTGACTTTGTGTCACCAGAGAACATCCATGAGTGTCTGCGTTTG
ACCGAAGAGTTTCGCCAACTTCCCAAAAACCACAAGGCCAGAGAGGACAAACTCGAGGCAAGTCTT
TTATCTCTTTGA

SEQ ID NO: 108, Proteína - *Arabidopsis thaliana*

MDSGVKLEHMNCFQLSYQYSWTTRKKRTLKPFMSKGSSPSSSSDSRKRKLSRAEDSDDSAVKRNAK
RRRKICKVEEYYEDDDCILSDWVQRNTAKRIDKRNEEVEVMVKIESGDDCTIGKWFSDVSSKRKDK
RQVEVDEDEEWEEEVTLCSKIKATSSRSRTHSLSANSPENVTDVISPCRSRSPASNVSDSIQKNDC
TSSRKQSGPICHQCLKGERITLLICSECEKTMFCLQCIRKWYPNLSEDDVVEKCPLCRQNCNCSKC
LHLNGLIETSKRELAKSERRHHLQYLITLMLPFLNKLSIFQKLEIEFEATVQGKLPSEVEITAAIS
YTDERVYCDHCATSIVDLHRSCPKCSYELCLKCCQEIREGSLSERPEMKFHYVDRGHRYMHGLDAA
EPSLSSTFEDEEANPSDAKWSLGENQSITCAPEKLGGCGERMLELRRILPLTWMSDLEHKAETFLS
SYNISPRMLNCRCSSLETELTRKSASRTTSSDNYLFCPESLGVLKEEELLHFQEHWAKGEPVIVRN
ALDNTPGLSWEPMVMWRALCENVNSTSSSEMSQVKAIDCLANCEVEINTRQFFEGYSKGRTYENFW
PEMLKLKDWPPSDKFEDLLPRHCDEFISALPFQEYSDPRTGILNIATKLPEGFIKPDLGPKTYIAY
GIPDELGRGDSVTKLHCDMSDAVNILTHTAEVTLSQEQISSVKALKQKHKLQNKVDKQSTEDCNEK
EEEEEEELNMPEISSNENEETGSALWDIFRREDVPKLEEYLRKHCKEFRHTYCSPVTKVYHPIHDQ
SCYLTLEHKRKLKAEYGIEPWTFVQKLGEAVFIPAGCPHQVRNLKSCTKVAVDFVSPENIHECLRL
TEEFRQLPKNHKAREDKLEASLLSL

SEQ ID NO: 109, DNA - *Arabidopsis thaliana*

ATGTATGGTAATGACTTGGATACTTCTGTTTACGGGAGTGGTTTTCCTCGAATAGGCGACCAAAGA
CCAGAATCAGTTGAGGCAGATATATGGGATGAGTATTGTGGTAGCCCCTGGAATCTCAATAACATG
CCTAAGTTGAAAGGATCTATGCTTCAGGCCATTCGACATAACATTAATGGTGTTACAGTGCCTTGG
CTATATCTTGGAATGCTCTTCTCTTCTTTTTGTTGGCATTTTGAGGACCATTGTTTTTACTCTGTC
AATTATCTACACTGGGGAGAGGCAAAATGTTGGTATGGTATTCCAGGCAGTGCTGCTAGTGCTTTT
GAAAAGGTCATGCGAAAAACCCTACCCGATCTCTTTGATGCTCAGCCAGATTTGCTCTTTCAACTA
GTTACCATGTTGAGTCCGACTGTTTTACAAGAAAATAAAGTCCCGGTCTACACAGTATTACAGGAG
CCAGGAAACTTCGTGATCACATTTCCAAAATCCTTTCATGCTGGATTCAATTTTGGTCTAAATTGT
GCAGAGGCCGTCAACTTTGCTACTGCTGATTGGCTACCTTATGGTGGTTCTGGGGCGGAGCTGTAT
CGGCTGTACCGGAAACCTTCAGTCATATCTCATGAAGAGCTTCTCTGCGTGGTAGCTAAGGGAAAC
TGCTGCAATAACGAAGGGTCAATACATTTGAAGAAAGAATTGCTCAGAATATACAGCAAGGAAAAA

FIGURA 10 CONT.

ACCTGGAGAGAGCAGCTCTGGAAAAGTGGTATTTTGAGATCTTCTCCTATGTTTGTGCCTGAATGC
GCTGATTCTGTGGGCATCGAAGAGGATCCAACATGCATCATATGCCAGCAGTTTCTTCATCTTTCT
GCTATCGTCTGCAATTGCAGGCCATCTGTTTTTGCATGCTTAGAGCACTGGAAGCATCTTTGTGAA
TGTGAACCTACGAAGTTGCGACTTGAATATCGTTATACCCTTGCCGAGTTAGATATGATGGTACAA
GAAGTTGAAAAGTTTGGTGGGTGCAAAACACAAGAAACCAAAATCTCACAACGGCCGAGTTCAGGC
ACCAAACGATCTATTGCTTTAAATAAAAAGCAGGAGGGAATGCAAGTGAGTCAGGCACGGCCAGCA
GACAAATGGCTTCTTAGAGCATCAAAGGTTCTTGACGCTGCGTTTTCTAGTGTTGAATATGCCACC
CTTTTGAAGGAATCAGAACAGTTTCTTTGGGCTGGATCAGAAATGGACCGTGTACGAGATGTTACA
AAAAGTTTGAACAAAGCAAAGATATGGGCTGAAGCTGTTAGTGACTGTCTTTCAAAAGTCGAAGGT
GAAGTCAACGATGATTCAATGAAAGTTCACTTGGAATTTATTGATGAGTTGCTGAGAGTTAATCCT
GTTCCTTGCTTTAATTCTGGTTATCTGAAACTAAAGGACTATGCTGAAGAAGCTAGAAAATTGTCC
GAGAAAATCGATTCAGCTCTCTCGAGTAGCCCAACTATCACCCAGCTGGAGCTATTGCATTCTGAA
GTCTCCAGATCGCCAATCTCCCTAAAAAAACACGAGATCTTGTCAAAGAAAATCTCTTCCGCAAAG
ATGTTAGCTAAAAGGGCGAAACGCTATCTCACAGACGCTAAACCTCCAGGAATTGAAATGGATGCA
CTTTTCAAGCTAAATTCAGAGATGTTAGAGCTTCATGTTCAACTTCCAGAAACAGAAGGGATTCTG
GATTTGGTAAAGAAATCAGAATCAGCCCGTGATAAAAGTAACAAAGTTTTGACTGGTTCTTTATCT
CTCGAGAATGTGGAAGAGTTGCTTCATGAATTCGATAGTTTCAGCATTAACGTTCCTGAGTTGAAT
ATCCTGAGGCAGTATCATGTTGACACTTTGTCTTGGATTTCACGCTTTAATGATGTAATGGTTGAT
GTTCGTGAAGGCAAGGACCAACGAAAGCTAATTAGTGACCTCAGTTCCCTTCTACGGGATGGAGCA
TCCTTAGGCATTCAAGTTGAAGGACTCCCTCTTGTTGAAGTTGAATTGAAGAAGGCATCTTGCCGG
GAAAAAGCTCGAACGGTTTATACTGCAAGAAAATCTCTGGATTTCATTGAGCAACTGCTCTCGGAA
GCTGTTATACTACACATCGAAGAGGAGGAGATATTTGTTGAGATCTCGGGAATTTTGTCAACAGCA
CGGTGTTGGGAGGAAAGAGCAAGCACTATCCTTGAAAATGAAACTCAGATGTACGAGCTTAAAGAT
CTCGTAAGGATGTCAGTCAACATTGATGCGGTTTTACCTACTCTGCAAGGCATAGAGAACACAATC
TCGTCGGCTGAAACTTGGCTTCAGAAATCTGAGCCTTTTCTATCAGCCACTTCCTCTATGGCATCA
TCTCCATGTTCTATGCTTGAACTTCCTGTATTAAAGGATCTAGTTACTCAGGCTAAATTGCTCAAT
GTTCAGCTTCAAGAGCCACGAATTCTTGAAACATTGTTGCTTAACTGCGAGAGGTGGCAGTGTGAT
AATCATCAGCTCTTGCAAGAAACTGAAGATTTGTTGGACAATGCGAAAATAGATGATGGCACGCAT
AGCAATATCCTTCCGAAGATAATGGATTTGATAACCAGAGTGGACTCCGCTAGGAGATCTGGTCTG
GCCCTTGGTCTCAATTTCGATGAACTTCCCAAACTTCGAACTGCAAGTCTAAAACTAGGATGGTGT
TGTAAGACCATCACGTTAAGCTCTAGCTCACCTACCTCTGAGTTGCTAGAAGATGTTGGAAAGCCC
TCGTTACAGCATATTCAGCAGCACTTAAAAGAGGGACAAACACTTGAAATATTACCTGAAGAGTAT
TACTTAGGCAAGAGACTCATGGAGTTAAAAGACACTGGACTGGAGTGGGCAAAACGAGCTAGAAAG
GTGGTAACAGACTCAGGTGCTCTTGCCTTGGAAGATGTTTTCGAGCTTATTTCTGAGGGTGAAAAC
TTACCTGTTCATGCAGAGCAGGAACTTCAGTCTTTACGAGCTCGGAGTATGTTACACTGCATTTGT
CTAAAGCCATACAACTCAAGATCCATGGTTTCTTGTAGTCAATGTGGCGAATGGTATCACACCTAT
TGTTTAAAGCTTCATTGGCGGCCTAAGGCTTATGTCTGCTCCGCTTGCTGTCCACTGGCAGAAACC
ACTCCACAGATCGATCCCGCCAGAGCAACAGAGCCAGAGAGACCGTCTCTGAACCAAAGACGGACA
AGAATGGTCGCGACTGATGCAGCAGTTAATGACTTAAAGTGGAAAACCCGAAAACACATCAAACGG
ACAACTAAACGGAGTCCTCAGGTTCATATTCTTCCCTGGTTTTTCACTTAA

SEQ ID NO: 110, Proteína - *Arabidopsis thaliana*

MYGNDLDTSVYGSGFPRIGDQRPESVEADIWDEYCGSPWNLNNMPKLKGSMLQAIRHNINGVTVPW
LYLGMLFSSFCWHFEDHCFYSVNYLHWGEAKCWYGIPGSAASAFEKVMRKTLPDLFDAQPDLLFQL
VTMLSPTVLQENKVPVYTVLQEPGNFVITFPKSFHAGFNFGLNCAEAVNFATADWLPYGGSGAELY
RLYRKPSVISHEELLCVVAKGNCCNNEGSIHLKKELLRIYSKEKTWREQLWKSGILRSSPMFVPEC
ADSVGIEEDPTCIICQQFLHLSAIVCNCRPSVFACLEHWKHLCECEPTKLRLEYRYTLAELDMMVQ
EVEKFGGCKTQETKISQRPSSGTKRSIALNKKQEGMQVSQARPADKWLLRASKVLDAAFSSVEYAT

FIGURA 10 CONT.

LLKESEQFLWAGSEMDRVRDVTKSLNKAKIWAEAVSDCLSKVEGEVNDDSMKVHLEFIDELLRVNP
VPCFNSGYLKLKDYAEEARKLSEKIDSALSSSPTITQLELLHSEVSRSPISLKKHEILSKKISSAK
MLAKRAKRYLTDAKPPGIEMDALFKLNSEMLELHVQLPETEGILDLVKKSESARDKSNKVLTGSLS
LENVEELLHEFDSFSINVPELNILRQYHVDTLSWISRFNDVMVDVREGKDQRKLISDLSSLLRDGA
SLGIQVEGLPLVEVELKKASCREKARTVYTARKSLDFIEQLLSEAVILHIEEEEIFVEISGILSTA
RCWEERASTILENETQMYELKDLVRMSVNIDAVLPTLQGIENTISSAETWLQKSEPFLSATSSMAS
SPCSMLELPVLKDLVTQAKLLNVQLQEPRILETLLLNCERWQCDNHQLLQETEDLLDNAKIDDGTH
SNILPKIMDLITRVDSARRSGLALGLNFDELPKLRTASLKLGWCCKTITLSSSSPTSELLEDVGKP
SLQHIQQHLKEGQTLEILPEEYYLGKRLMELKDTGLEWAKRARKVVTDSGALALEDVFELISEGEN
LPVHAEQELQSLRARSMLHCICLKPYNSRSMVSCSQCGEWYHTYCLKLHWRPKAYVCSACCPLAET
TPQIDPARATEPERPSLNQRRTRMVATDAAVNDLKWKTRKHIKRTTKRSPQVHILPWFFT

SEQ ID NO: 111, DNA - *Arabidopsis thaliana*

ATGACGACGTTAGGACAAAGAGATCGCCGACCAGATGCTCTCGGTAGTCTCAGTGTTCTACCAGAT
GAGACCATCTGTGTTCTTCTCGAATACCTTGCTCCCCGAGATATTGCTCACCTCGCTTGTGTCAGC
AGTGTGATGTATATTCTATGTAATGAAGAACCATTGTGGATGAGTTTGTGTCTCAGAAGAGCAAAA
GGTCCTCTTGAATACAAAGGTTCTTGGAAAAAAACAACATTGCATCTAGAAGGAGTTACTCAAGAA
AATGACGCATACAGAAAATGTTTTCATTTTGATGGATTCATGTCGTTGTACTTGTATAAACGATTC
TATAGGTGTAATACATCTCTTGACGGGTTCTCTTTTGATAATGGGAATGTGGAACGTAGGAGAAAT
ATCTCCTTGGATGAGTTTTCTAAGGAATACGACGCCAAGAAACCTGTTTTGCTTTCTGGCCTTGCT
GATTCTTGGCCAGCCAGTAATACGTGGACTATCGACCAACTCTCAGAGAAATATGGTGAAGTCCCG
TTTAGAATATCTCAGAGGAGCCCCAACAAAATTTCCATGAAGTTCAAGGATTACATCGCATATATG
AAAACTCAGCGGGATGAAGATCCTCTATACGTTTTCGATGACAAGTTTGGAGAAGCTGCTCCAGAA
TTATTGAAAGACTATAGTGTGCCCCATTTGTTTCAAGAAGATTGGTTTGAGATCTTAGACAAGGAA
AGTCGTCCTCCATACAGATGGCTTATAGTTGGTCCGGAGAGGTCTGGTGCATCTTGGCATGTTGAC
CCAGCTCTTACCAGCGCCTGGAACACCCTGCTTTGCGGTCGGAAAAGGTGGGCATTGTATCCCCCT
GGAAAAGTGCCTCTAGGTGTTACAGTCCATGTCAATGAAGATGAGGGTGATGTCAGCATTGATACA
CCCTCATCTCTGCAGTGGTGGCTAGACTATTATCCCCTTCTTGCGGACGAAGACAAACCGATTGAG
TGCACACTACTACCTGGCGAAACGATTTATGTTCCAAGTGGTTGGTGGCACTGTATCCTTAATCTT
GAACCAACAGTGGCTGTTACCCAGAATTTTGTGAACAAAGAAAACTTTGGGTTCGTGTGCTTGGAT
ATGGCGCCTGGTTACCATCACAAAGGAGTTTGCCGTGCTGGGCTTCTTGCTCTTGATGATGAAAAT
TCTGAAGATTTGGAAGAAGAAACACACGATGAAGAAGATAATACTTTGAGCTATTCAGACCTTACC
AGGAAAGAGAAGAGGACACGGATGAATGGAGGTGGAGAGACTGAAAACCGCGAAGAGGATGTGAAT
GGAGTATCGAAGAGATATAATATGTGGAAGAATGGATTTTCATATGATATCGATTTCCTCGCTTCA
TTTCTTGATAAAGAAAGAGATCATTACAATTTCCCATGGTCAATGGGGAACTCTGTAGGCCAACGA
GAAATGAGAGCCTGGCTATCCAAGCTTTGGGTTCTGAAGCCTGAGATGAGAGAACTAATATGGAAG
GGAGCATGCATTGCTTTAAATGCTGAGAAATGGTTGCGATGCCTAGAGGAAGTATGTACTTTCCAC
AACTTGCCGTTGGTTACCGAAGATGAAAAGCTTCCAGTTGGAACTGGCAGCAACCCTCTTGAGTTT
TATGACATTCTTGGCCGGGCTGATTCTCCTTTGAAAACACATATTCCTGAAGTTCTAGCAAGCGGG
ATTCTTTTCTTCGAAAAAGGATCTTACAAAGTTGTCCCTTGGGATGGCAAGAGAATCCCAGATATT
ATCTCTAGCTCCAGTTTCGATTTTGATGCATCCATGTTAAACAGCGAATTCCCATTTGGTATTTGG
AATAAAACGCTACGTGAACATAAAAACCAGGGGAAGCCAGCACCCGATTCTTTTGGTTCATTGAGT
TCACATGTTTGGCCATATATTATAACCAAAAGATGCAAAGGGAAGATTTTTGCCCAACTAAGAGAT
GATCTAACCTGGAACGACGCTCAGAACCTGGCTTTCTTTCTCGGACAACAACTACGCAACCTTCAT
CTACTTCCATATCCACCGGTCACACGTCCGGAGTTGTTGAATGTGAATGCTGTTCATGAGGAGTTG
AACATTCCAGCAGAATGGAAAGTTTTTGTCGATGCTCTGTGCCAAAAGAAGAAGGACGTCACTAGT
CGTCTGGAAAACTGGGGAAATCCGATACCTCGGGCTCTGATGACCAAGATAGACGAATACATTCCC
GATGACTTTTTTGTAGATTTACTCCACGTCTTCAAGGAGACAAATGGTGGAGATGAAATCAAGCCC

FIGURA 10 CONT.

TGCACCTGGATACACTCAGACGTAATGGATGACAACATTCACATGGAACCATACGCAGATGATTCA
GTTGATGGTCAACACAACTCATGGCGTCCCAGTCATATTCTTGACTTCAGCGATTTAACCATAGGA
GATCCCATCTGCGACTTGATACCAATATACTTGGACGTCTTTAGAGGAGACGCTGATCTTCTTAAG
AAGCTTCTAGAAAATTATGGACTTCCATTAATCAGAAGTAGATCTTCAGAGAACGGAACAACAAAA
ACAGCAGACAGTACGAGGAAGAAAGTATTGTCTCCATCCTACCGTACAATGTGTTACTGTATATTA
CATGAAGAAAATGTGTTGGGTTCAATATTCAGTATTTGGGACGAACTTCGGACTGCTGAATCATGG
GAACAAGTTGAGCAAACTGTTTGGAGTCTACTTAACACCTACTAA

SEQ ID NO: 112, Proteína - *Arabidopsis thaliana*

MTTLGQRDRRPDALGSLSVLPDETICVLLEYLAPRDIAHLACVSSVMYILCNEEPLWMSLCLRRAK
GPLEYKGSWKKTTLHLEGVTQENDAYRKCFHFDGFMSLYLYKRFYRCNTSLDGFSFDNGNVERRRN
ISLDEFSKEYDAKKPVLLSGLADSWPASNTWTIDQLSEKYGEVPFRISQRSPNKISMKFKDYIAYM
KTQRDEDPLYVFDDKFGEAAPELLKDYSVPHLFQEDWFEILDKESRPPYRWLIVGPERSGASWHVD
PALTSAWNTLLCGRKRWALYPPGKVPLGVTVHVNEDEGDVSIDTPSSLQWWLDYYPLLADEDKPIE
CTLLPGETIYVPSGWWHCILNLEPTVAVTQNFVNKENFGFVCLDMAPGYHHKGVCRAGLLALDDEN
SEDLEEETHDEEDNTLSYSDLTRKEKRTRMNGGGETENREEDVNGVSKRYNMWKNGFSYDIDFLAS
FLDKERDHYNFPWSMGNSVGQREMRAWLSKLWVLKPEMRELIWKGACIALNAEKWLRCLEEVCTFH
NLPLVTEDEKLPVGTGSNPLEFYDILGRADSPLKTHIPEVLASGILFFEKGSYKVVPWDGKRIPDI
ISSSSFDFDASMLNSEFPFGIWNKTLREHKNQGKPAPDSFGSLSSHVWPYIITKRCKGKIFAQLRD
DLTWNDAQNLAFFLGQQLRNLHLLPYPPVTRPELLNVNAVHEELNIPAEWKVFVDALCQKKKDVTS
RLENWGNPIPRALMTKIDEYIPDDFFVDLLHVFKETNGGDEIKPCTWIHSDVMDDNIHMEPYADDS
VDGQHNSWRPSHILDFSDLTIGDPICDLIPIYLDVFRGDADLLKKLLENYGLPLIRSRSSENGTTK
TADSTRKKVLSPSYRTMCYCILHEENVLGSIFSIWDELRTAESWEQVEQTVWSLLNTY

SEQ ID NO: 113, DNA - *Arabidopsis thaliana*

ATGGAGCCTTTCAGTGCTGCTCAGAACAAAGAGGACAAGGACACTAGCGTAGAACCTCCTCGTAGG
CGTTGTCATCGAAAGAACAAGGGCACTAATGTAGAACCTCCTAGTAGTCCGTATCATCCGAAGGTT
CTTGCTAGATGGGATCCGGCTAATGAAAAGAGGCCAGACATTGGTGAAGCTCCAGTATTCCACCCT
ACATCAGAGGAGTTTGAAGATACGCTTGCTTACATAGAAAAAATACGTCCCTTAGCTGAATCATTT
GGTATTTGTCGAATCGTGCCACCGTCGAATTGGTCTCCTCCTTGCCGGCTTAAAGGGGATAGTATA
TGGAAGAATAAAAATTTCCCAACCCGTGTTCAGTTTGTTGACCTGCTTCAGAACAGAGGACCCGTG
AAGAAGAAGACACCTAAAGGAAGGAAACGGAAACGAGGAAAATACTCAAGAACCGTCGCCCCCAAG
AAACGAAATGGCTCTGTTTCTAAATCAGTTTCTACCCCCAAGGCGACCGAAGAAGAGAATTTTGGT
TTCGAATCTGGTCCAGAATTTACTCTTGAGAAGTTTGAGAAATACGCTCAAGATTTCAAAGACTCC
TACTTTGAAAGAAAAGACAATGTTGGAGATCCATCTGTTGAGGAGATCGAAGGAGAGTACTGGAGA
ATCATTGAGAAAGAAACAAATGAAGTTAAGGTATTGTATGGGACGGATTTGGAGAATCCCATACTT
GGAAGCGGTTTCTCCAAAGGGGTAAAAATTCCAACCAGAAGAAATGATATGGATAAATACATATCT
TCTGGTTGGAATTTGAACAACCTCGCCCGTCTGCAGGGATCTCTACTATCCTTTGAGGATTGTGAG
ATCTCAGGGGTTCAGGTGCCATGGCTCTATGTGGGCATGTGCTTTTCAACGTTTTGTTGGCATGTT
GAGGACAATCATCTATATTCACTGAACTATCATCACTTTGGCGAGCCAAAAGTATGGTACGGAGTT
CCGGGAAGCCATGCAACTGGGCTAGAGAAGGCAATGAGAAAACATCTACCAGATTTGTTCGATGAA
CAACCTGATCTACTTCATGAACTGGTTACTCAGTTTTCTCCAACGATTTTGAAAAATGAGGGAGTC
CCGGTTTATCGAGCTGTTCAGAATGCAGGGGAGTATGTTCTAACATTCCCTAGAGCATATCATTCG
GGTTTCAACTGCGGATTCAACTGTGCAGAAGCGGTGAATGTGGCTCCGGTTGATTGGCTGGCTCAT
GGACAGAATGCTGTGGAAATATATAGCCAAGAGACTCGGAAAACGTCTCTATCTCATGACAAAATT
CTCCTTGGAGCAGCTTTTGAAGCTGTTAAGTCCCTCTCAGCACATGGGGAAGATAATACAAAAAGG
TTCAGCTGGAAGAGATTCTGTGGGAAGGACGGTATCATAACTAAGGCCATCGAGGCACGGTTACGG

FIGURA 10 CONT.

ATGGAAGAAAAAAGAATTGAGGCTCTTGGAAATGGTTTCAGCTTGGTAAAGATGGATAAAGACTTT
GATTCAAACTGTGAAAGGGAATGCATCTCTTGCTTTAGTGACTTGCATCTCTCTGCTACTGGCTGC
AAGAATTGCTCATCTCTCGAAGAATATGGTTGCACGAAACACGATATTTGTTCATGTGAAGGGAAA
GACCGTTTCATTTTTCTTCGTTACACCATAGATGAGTTAAGCTCATTGGTAAGAGCATTGGAAGGT
GAATCAGATGATCTAAAAGCTTGGCTTTCCAAAGTCATGGAAGGTTGTTCAGAGACTCAGAAGGGA
GAATCTAGTGGGATTATCGTTAAGGAAAAGCAGGTACAAGAAGAATGTTTTGATCTGAATGGTGAG
TGTAATAAATCTTCTGAGATATGTGAGGATGCATCCATCATGGACTTAGCTGCTTATCATGTTGAA
CCTATAAATCTTGGGTTCTTGGTCGTGGGAAAGCTTTGGTGTAACAAGCATGCTATATTTCCAAAA
GGGTTTAAGAGTCGTGTTAAGTTCTATAACGTGCAAGATCCAATGAGAATAAGCTATTACGTCTCT
GAGATTGTAGATGCAGGACTTCTGGGTCCACTCTTCAAGGTTACTTTGGAAGAATCTCAAGACGAG
AGCTTCTCTTATGCCTCACCTCAGAAGTGCTGGGAAATGGTGTTGCTTAGAGTTAAGGAAGAGATC
ATGAGACGCAGCAACCAAAAACAAGATGTCCATATGTTGGAAAGCATCGATGGGTTAAAAATGTTT
GGCTTTCGCTCTCCGTTTATTGTTCAGGCCACTGAGGCTCTTGACCCGAACCATGGCCAAGTGGAG
TATTGGAACCATAAGAATGAGAAGGATTCGCTGGAAATGAAAGATTGCTTCATGTCAAATTCTTCC
CAGAGCTTAAGCAAAGCAAGACTTTTTGGGGTTGATTTGAATTGA

SEQ ID NO: 114, Proteína - *Arabidopsis thaliana*

MEPFSAAQNKEDKDTSVEPPRRRCHRKNKGTNVEPPSSPYHPKVLARWDPANEKRPDIGEAPVFHP
TSEEFEDTLAYIEKIRPLAESFGICRIVPPSNWSPPCRLKGDSIWKNKNFPTRVQFVDLLQNRGPV
KKKTPKGRKRKRGKYSRTVAPKKRNGSVSKSVSTPKATEEENFGFESGPEFTLEKFEKYAQDFKDS
YFERKDNVGDPSVEEIEGEYWRIIEKETNEVKVLYGTDLENPILGSGFSKGVKIPTRRNDMDKYIS
SGWNLNNLARLQGSLLSFEDCEISGVQVPWLYVGMCFSTFCWHVEDNHLYSLNYHHFGEPKVWYGV
PGSHATGLEKAMRKHLPDLFDEQPDLLHELVTQFSPTILKNEGVPVYRAVQNAGEYVLTFPRAYHS
GFNCGFNCAEAVNVAPVDWLAHGQNAVEIYSQETRKTSLSHDKILLGAAFEAVKSLSAHGEDNTKR
FSWKRFCGKDGIITKAIEARLRMEEKRIEALGNGFSLVKMDKDFDSNCERECISCFSDLHLSATGC
KNCSSLEEYGCTKHDICSCEGKDRFIFLRYTIDELSSLVRALEGESDDLKAWLSKVMEGCSETQKG
ESSGIIVKEKQVQEECFDLNGECNKSSEICEDASIMDLAAYHVEPINLGFLVVGKLWCNKHAIFPK
GFKSRVKFYNVQDPMRISYYVSEIVDAGLLGPLFKVTLEESQDESFSYASPQKCWEMVLLRVKEEI
MRRSNQKQDVHMLESIDGLKMFGFRSPFIVQATEALDPNHGQVEYWNHKNEKDSLEMKDCFMSNSS
QSLSKARLFGVDLN

SEQ ID NO: 115, DNA - *Arabidopsis thaliana*

ATGGGCATTGAAGGTGTATCAACGTATTTGAAATCTGGAAATATGGACACAATTTCTGCTCCGCCA
GGTTTTGTATCTCAAACATCGTTTGTGCTGAGGAATGTACCCCGAGACAAAGAAAGTCCCAGGTCC
GTGTCGAGACAAGAGCAGACAACTGGATTCGGCACAGATGATAAAGATAGTTGCAATATGTTTCTT
AAAAGCCGACCGTGGATAGTGCATGGTCACACAATTCCCAGCTCAGAGGCCTTAAGGCCGAAAAAA
ACTGAGGTACGGAGAAGACGGCCCCTCAAAGTGTCTGAGACCAAAGTCCTTGAGGAAGCTCCTGTG
TTCAACCCAACCGAAGAGGAATTCAGGGACACACTCTCATATATATCAAGCTTGCGTGACAGAGCT
GAGCCGTATGGAATCTGTTGTGTTGTTCCTCCACCTTCATGGAAACCTCCATGTCTTCTCAAGGAG
AAACAAATATGGGAGGCTTCTACATTTTTCCCTCAAGTTCAGCTCTTTGGAATTCAGACCGAGAAC
CGAAAGATTAAAAAGGAGGTTGATGCAGATAGTAATGATGCTGCATCTGAAGGGGTTCAGTTATGT
AGAGTAGAACGTGGTCCTGGGTATACTTTGAAGTCCTTTAAAAACTTTGCAGATACATACAAGAAG
AGTCATTTTGGCATGAAGGATGAGGTTTTGGGTTCAGAGAACTCATCTCCATCACTTAAACCAAAT
GAGCTAATTGTGGCAGACATTGAAAAAGAGTACAGACAAATTGTTGAGAGTCCACTTATTGAAATT
GGGGTGCTTTATGGCAATGATTTAGACACTGCAACGTTTGGAAGTGGATTTCCCTTATCTGCACCA
TCTGAATCTAGCAAATATTCATCAGGTTGGAATCTGAACAGTACGGCTAAGCTTCCTGGTTCTCTT
CTATCATTAGAAGACTGCGAATCAGTTTGTGTTCCTCGTCTAAGTGTAGGAATGTGCTTGTCTTCC
CAGTTCTGGAAATCTGAGAAGGAGAGACTCTACTCTCTCTGCTATTTGCATGTGGGCGCTCCTAGA

FIGURA 10 CONT.

GTGTGGTATTCCGTAGCAGGATGTCATCGATCAAAGTTTAAGGCTGCCATGAAAAGCTTCATCCTT
GAGATGTCAGGAGAACAGCCAAAGAAGAGTCATAATCCTGTGATGATGATGTCTCCATATCAATTG
AGCGTGGAGGGTATACCAGTGACCCGCTGTGTCCAGCACCCTGGCCAGTATGTTATTATATTTCCA
GGGTCGTATTACTCTGCCTTTGACTGTGGCTTCAACTGTTTGGAGAAAGCAAACTTTGCTCCCCTT
GATTGGTTGCCCCATGGGGATATTGCTGTTCAGGTGAATCAAGAGATGAGTAAAACGTCGCTGATA
TCGTATGATAAGCTATTATTTAGTGCTGCAAGGGAAGCTGTAAAATGTCTTAAGGAATATGGGTTG
TCGAAGAAAAACACAGCATGTTACACGAGATGGAATGATTCTTGTGGGACGGATGGGTTGTTCTCC
AACATAATCAAGTCACGGATCAAACTGGAGAAAAATAGACGTGAATTTCTTATCAGTTCACTAGAA
TCGCAAAGGATGGACAAGAGTTATGATGCTGTGAACAAAAGGGAATGCTGTGTATGTCTTGGGGAT
TTGTATCTTTCCGCGGTTAACTGTTCATGCTCTGCTAATCGTTACTCGTGTCTGAACCACATGAGG
AAGCTCTGTGCATGTCCGTGTGACAGAAAGAGTTTTCTCTATAGGTACACTATGGATGAGTTGAAT
CTTCTCGTTGAAGCCCTGGAAGGTAAAAAGCTCAGCTCTATGTTCAGATGGGCAGGCATAGATCAA
AAATTCTGTGCTTCTCCAGCCACCACAAGTTCAAAACCCGAAGAAGACAAAGGCAAGGAAACAGAT
GAGGTTACACCCTGTAATATCACGAGGAAAGATGTCGCAGCTGGAACTAAGGACCAGACAAGAGTG
AAGGCGAGATCTTTGGCTGATATACTGAATGTTAAAGATGGAAACAATGATGCAAAGGAGACTTTA
GAATCTTGTTCTAAGAAATCAAACAGGCCATGCGATAACGATTCCTCTGAGGCCAACGCGCCAAAG
AAACAGAAGCAGTGA

SEQ ID NO: 116, Proteína - *Arabidopsis thaliana*

MGIEGVSTYLKSGNMDTISAPPGFVSQTSFVLRNVPRDKESPRSVSRQEQTTGFGTDDKDSCNMFL
KSRPWIVHGHTIPSSEALRPKKTEVRRRRPLKVSETKVLEEAPVFNPTEEEFRDTLSYISSLRDRA
EPYGICCVVPPPSWKPPCLLKEKQIWEASTFFPQVQLFGIQTENRKIKKEVDADSNDAASEGVQLC
RVERGPGYTLKSFKNFADTYKKSHFGMKDEVLGSENSSPSLKPNELIVADIEKEYRQIVESPLIEI
GVLYGNDLDTATFGSGFPLSAPSESSKYSSGWNLNSTAKLPGSLLSLEDCESVCVPRLSVGMCLSS
QFWKSEKERLYSLCYLHVGAPRVWYSVAGCHRSKFKAAMKSFILEMSGEQPKKSHNPVMMMSPYQL
SVEGIPVTRCVQHPGQYVIIFPGSYYSAFDCGFNCLEKANFAPLDWLPHGDIAVQVNQEMSKTSLI
SYDKLLFSAAREAVKCLKEYGLSKKNTACYTRWNDSCGTDGLFSNIIKSRIKLEKNRREFLISSLE
SQRMDKSYDAVNKRECCVCLGDLYLSAVNCSCSANRYSCLNHMRKLCACPCDRKSFLYRYTMDELN
LLVEALEGKKLSSMFRWAGIDQKFCASPATTSSKPEEDKGKETDEVTPCNITRKDVAAGTKDQTRV
KARSLADILNVKDGNNDAKETLESCSKKSNRPCDNDSSEANAPKKQKQ

SEQ ID NO: 117, DNA - *Arabidopsis thaliana*

ATGGATTCTGTGGAGGAAGAAGGTGTTGTTAGGGTTGAAGAAGAGAACGGACGCGGTGGTCTCCGT
AGGCATCGGAGGGTTTCTACGAAGCTGGCGAATTACGTAGATCCTCCGACTGACGACGAAGAAGAT
GGAGGACCCAAGCGAAAGGGCAAAAGGGGTGGAAATAGGGCTCCGAAAAAGACACCGAAGAAAGAC
GAGGAGATGCAAAAGAACGAAATTGATGAAGCAAACCGAGTCACTGGTCTGGTGAAGGAAAAAAGA
GCCGCGACTAAGATTCTGAATCGCAAAGACTCCATTATTGAAGTAGGAGAAGCTTCTGGAAGCATG
CCCAAGGAAGTAAAGGGAATTAGGATTGGGAAAAGGAAGGGAGAAATCGATGGTGAGATTCCAACT
AAGCCGGGCAAGAAGCCAAAGACTACAGTAGACCCGAGAATCATTGGTTACAGACCGGACAATATG
TGCCATCAATGTCAAAAGAGCGATAGGATTGTTGAACGGTGTCAGACCTGCAATAGTAAGCGTTAT
TGTCATCCATGCTTGGACACTTGGTACCCCCTTATAGCAAAAGAAGATGTTGCCAAGAAATGCATG
TTCTGCTCTAGTACTTGCAATTGCAGAGCATGCTTGCGTCTAGATACTAAATTGAAAGGGATAAAT
TCGAACCTCATAGTCAGCGAGGAAGAAAAGGTCCAAGCCTCTAAGTTTATTCTGCAGAGCCTTCTC
CCACATCTGAAAGGAATAAATGATGAACAAGTTGCTGAGAAGGAAGTTGAGGCCAAAATATACGGA
CTGAAGTTTGAAGAGGTGAGGCCCCAAGACGCTAAAGCTTTCCCTGATGAAAGACTATACTGTGAT
ATCTGCAAGACTTCTATCTATGATCTCCATAGGAATTGCAAGTCCTGTAGTTTTGATATCTGCCTT
AGCTGTTGCCTTGAGATCCGCAATGGAAAGGCTCTGGCATGTAAGGAGGATGTGTCTTGGAATTAT

FIGURA 10 CONT.

ATTAACCGAGGTTTAGAGTATGAACATGGACAAGAAGGAAAAGTGATTGAAAAGCCGGCGAATAAG
CTAGATGATAAGTTGAAAGATAAGCTGGATGGTAAGCCGGATGATAAGCCGAAGGGTAAGCCAAAG
GGTAGGCCAAAGGGTAAGCCGGATGATAAGCCGAAGGGTAAACTGAAGGGTAAGCAGGATGATAAA
CCGGATGATAAGCCAGATGAGAAGCCGGTTAATACAGACCATATGAAGTACCCTTCTCTGTGGAAA
GCAAATGAAGCTGGGATCATTACTTGTTGTTGTGGTGCTGGGGAGTTAGTGCTGAAACGACTACTT
CCGGATGGTTGGATATCTGAGTTGGTCAACAGAGTTGAAAAAACTGCTGAAGCTGGCGAGCTTTTG
AATTTACCTGAAACGGTTTTGGAGCGATGCCCTTGTTCTAACTCTGACAGACATATTGACATAGAC
AGCTGTAATTTGTTAAAAGCTGCTTGTCGAGAAGGTTCAGAAGACAATTATCTTTACTCTCCAAGT
GTATGGGATGTTCAACAAGATGATTTGAAGCATTTTCAGCACCATTGGGTAAAAGGCGAGCCTGTG
ATTGTGAGAAATGTGCTTGAAGCTACATCTGGTTTAAGCTGGGAACCAATGGTAATGCATCGTGCC
TGCCGCCAGATAAGCCATGTCCAGCATGGATCACTTAAGGATGTTGTTGCTGTTGATTGTTTGGAC
TTCTGTGAGGTAAAAGTTAATCTTCATGAATTTTTTACTGGATACACGGATGGCCGATATGATCGA
ATGGGTTGGCCACTAGTTCTGAAACTGAAAGATTGGCCTCCAGCTAAAGTCTTCAAAGATAACTTG
CCACGTCATGCTGAGGAGTTCTTATGTAGTTTGCCTTTGAAGCATTACACTCATCCGGTTAATGGA
CCTTTGAATCTCGCTGTCAAGCTTCCCCAAAATTGCTTGAAGCCAGACATGGGACCAAAGACGTAC
GTTGCTTCTGGATTTGCTCAAGAATTAGGCCGTGGAGATTCTGTTACTAAGCTCCACTGCGACATG
TCTGATGCGGTGAATATTTTGACGCACATATCTGAAGTGCCCAATATGCAACCTGGAATTGGAAAT
CTGAAAAAGAAGCATGCCGAACAAGATCTCAAGGAGCTGTATAGCTCAGTAGCTAACAAGGAGGAA
ATGATGGAGATTCTTGAAAATTCCAGACAACAAGTCCAAAATGTTGAAACTGATGATGGAGCTCTT
TGGGACATTTTCCGTAGAGAAGATATTCCTAAATTAGAAAGTTACATTGAGAAACATCACAAGGAG
TTCAGACATCTCTACTGCTGTCCTGTGTCTCAGGTTGTTCATCCTATACATGACCAGAACTTCTAT
CTGACGCGGTATCATATAATGAAGCTGAAAGAAGAATATGGCATTGAACCTTGGACCTTCAATCAG
AAGCTTGGTGATGCTGTTTTGATACCTGTAGGTTGCCCTCATCAAGTTAGAAATTTGAAGTCCTGC
AATAAGGTAGCGCTTGACTTCGTCTCACCTGAAAATGTCAGCGAGTGTTTACGCTTGACAAAACAG
TATCGTCTACTTCCACCAAATCATTTTGCAAAAGAAGACAAGTTAGGGGTTAAGAAGATGATAGTC
CATGCAGTTGATAAAGCTCTCAGAGACTTAAGTGGAGAGAAGTCTCCAGAACCTGAAGAGAAGAAA
CAAAATATGAGAGGGCCGAAGAAGGGGGCAGCCAAGGCAGTTGCCAAGGCTCTCAAAGACTTATCT
CCAAGTGAAAAGAAGTCTTCGGAAGCAGCTGAAGAGGAGATATCAAATGGGATAGTCAATGCAATC
GATAAGGGTCTCAAAGACTTACCTCCAAGTGAAGAGAAGTCTTCCGAAGCCAAAGTGGAGATATCA
AATGGGATAGTCAGTGCAATGGATAAGGATCTCGAACACATATCTTCAAGCGAAAAGAAGTCTACG
GAAGAGGAAGGGGTGAAAAGACCAAATATTGTGAGGACATATGAGCGGAGAAAAAAGCTGGGAAGT
GAAGTAACCAACGCATACATTGACAGATTAGAAATGGAGAAGATGTAG

SEQ ID NO: 118, Proteína - *Arabidopsis thaliana*

MDSVEEEGVVRVEEENGRGGLRRHRRVSTKLANYVDPPTDDEEDGGPKRKGKRGGNRAPKKTPKKD
EEMQKNEIDEANRVTGLVKEKRAATKILNRKDSIIEVGEASGSMPKEVKGIRIGKRKGEIDGEIPT
KPGKKPKTTVDPRIIGYRPDNMCHQCQKSDRIVERCQTCNSKRYCHPCLDTWYPLIAKEDVAKKCM
FCSSTCNCRACLRLDTKLKGINSNLIVSEEEKVQASKFILQSLLPHLKGINDEQVAEKEVEAKIYG
LKFEEVRPQDAKAFPDERLYCDICKTSIYDLHRNCKSCSFDICLSCCLEIRNGKALACKEDVSWNY
INRGLEYEHGQEGKVIEKPANKLDDKLKDKLDGKPDDKPKGKPKGRPKGKPDDKPKGKLKGKQDDK
PDDKPDEKPVNTDHMKYPSLWKANEAGIITCCCGAGELVLKRLLPDGWISELVNRVEKTAEAGELL
NLPETVLERCPCSNSDRHIDIDSCNLLKAACREGSEDNYLYSPSVWDVQQDDLKHFQHHWVKGEPV
IVRNVLEATSGLSWEPMVMHRACRQISHVQHGSLKDVVAVDCLDFCEVKVNLHEFFTGYTDGRYDR
MGWPLVLKLKDWPPAKVFKDNLPRHAEEFLCSLPLKHYTHPVNGPLNLAVKLPQNCLKPDMGPKTY
VASGFAQELGRGDSVTKLHCDMSDAVNILTHISEVPNMQPGIGNLKKKHAEQDLKELYSSVANKEE
MMEILENSRQQVQNVETDDGALWDIFRREDIPKLESYIEKHHKEFRHLYCCPVSQVVHPIHDQNFY
LTRYHIMKLKEEYGIEPWTFNQKLGDAVLIPVGCPHQVRNLKSCNKVALDFVSPENVSECLRLTKQ
YRLLPPNHFAKEDKLGVKKMIVHAVDKALRDLSGEKSPEPEEKKQNMRGPKKGAAKAVAKALKDLS
PSEKKSSEAAEEEISNGIVNAIDKGLKDLPPSEEKSSEAKVEISNGIVSAMDKDLEHISSSEKKST
EEEGVKRPNIVRTYERRKKLGSEVTNAYIDRLEMEKM

FIGURA 10 CONT.

SEQ ID NO: 119, DNA - *Arabidopsis thaliana*

ATGGCGGTTTCAGAGCAGAGTCAAGATGTGTTTCCATGGCTTAAATCGTTACCGGTTGCTCCAGAG
TTCAGACCTACTCTAGCAGAGTTTCAAGATCCGATCGCTTACATTTTGAAGATTGAAGAAGAAGCT
TCTAGATATGGAATCTGTAAAATTCTGCCTCCACTGCCTCCTCCTTCCAAAAAAACCTCGATTTCT
AATCTCAACCGTTCTCTGGCGGCAAGAGCGGCGGCGAGGGTTCGTGACGGCGGCTTTGGCGCGTGT
GATTATGACGGTGGTCCCACATTCGCCACGCGCCAGCAGCAGATCGGGTTTTGCCCTAGGAAACAG
CGTCCAGTGCAGAGACCTGTGTGGCAGAGTGGAGAGGAGTACTCTTTTGGTGAGTTCGAGTTTAAA
GCTAAGAACTTTGAGAAGAATTATCTCAAGAAATGTGGTAAAAAGAGTCAACTCTCTGCCCTTGAA
ATTGAAACACTTTACTGGAGAGCCACTGTGGATAAACCCTTCTCTGTTGAGTATGCCAATGACATG
CCTGGCTCGGCTTTTATCCCTCTGAGTCTGGCTGCTGCGAGGAGGAGAGAGTCTGGTGGTGAAGGA
GGAACAGTTGGTGAGACCGCTTGGAACATGAGGGCAATGTCTAGAGCCGAAGGATCATTGCTTAAG
TTCATGAAGGAAGAGATCCCTGGAGTTACATCACCAATGGTGTATGTTGCTATGATGTTTAGTTGG
TTTGCTTGGCATGTGGAGGACCATGACCTTCATAGTCTCAATTACTTGCATATGGGTGCTGGTAAG
ACTTGGTACGGTGTGCCAAAGGATGCTGCTCTGGCTTTTGAGGAGGTTGTTAGGGTTCATGGTTAC
GGTGAAGAGCTCAATCCTCTTGTGACATTTTCTACTCTTGGTGAGAAGACAACTGTGATGTCTCCT
GAAGTATTTGTTAAAGCCGGAATACCGTGTTGCAGGTTAGTGCAAAATCCTGGAGAGTTTGTCGTC
ACCTTTCCGGGAGCTTATCATTCGGGATTTAGTCATGGATTTAATTTTGGAGAAGCATCTAACATT
GCCACTCCCGAATGGTTGAGAATGGCTAAAGATGCTGCTATCCGGCGAGCTGCTATAAATTACCCT
CCAATGGTTTCTCATCTCCAGCTACTTTATGACTTTGTATTGGCTCTAGGTTCTAGAGTGCCAACA
AGCATCAATCCCAAACCACGGAGTTCTAGATTAAAAGATAAGGCAAGAAGCGAAGGAGAAAGATTG
ACCAAAAAGCTATTTGTGCAAAACATTATCCACAACAACGAATTGCTTTCTTCTCTCGGAAAAGGA
TCCCCAGTGGCCCTTCTCCCACAGAGTTCCTCAGATATATCAGTTTGTTCTGACCTGCGAATTGGA
TCCCATTTGATAACCAACCAGGAAAACCCAATCCAGTTAAAGTGTGAGGACTTAAGTTCTGATAGT
GTTGTGGTTGATCTCAGTAACGGTTTAAAGGATACAGTTTCAGTGAAAGAAAAATTTACATCTTTA
TGTGAAAGGAGCAGAAATCACCTAGCAAGCACGGAGAAGGACACTCAAGAAACTCTGTCTGATGCT
GAAAGGAGGAAGAATGATGCAGCTGTTGCGCTTTCGGATCAAAGGCTTTTCTCTTGTGTTACATGT
GGAGTCTTAAGCTTTGATTGTGTAGCTATCGTTCAACCTAAAGAAGCAGCTGCTAGATATCTCATG
TCTGCAGATTGTAGCTTCTTCAATGATTGGACAGCTGCTTCTGGATCTGCAAATCTTGGTCAGGCT
GCAAGATCACTTCATCCTCAAAGCAAAGAGAAGCATGATGTAAATTACTTCTACAATGTTCCTGTT
CAAACTATGGATCATTCAGTGAAGACTGGCGATCAAAAAACTTCAACAACTTCCCCGACAATAGCG
CATAAAGATAATGATGTTCTTGGGATGTTAGCTTCAGCATATGGAGACTCTTCTGATTCCGAGGAA
GAAGATCAAAAAGGCTTAGTTACCCCTAGTTCCAAAGGGGAAACAAAAACGTATGATCAAGAAGGT
TCAGATGGCCATGAGGAGGCTAGAGATGGTAGAACTTCTGATTTTAACTGCCAGAGACTAACCAGC
GAACAGAATGGGTTAAGCAAAGGCGGAAAATCATCACTTCTGGAAATAGCTTTACCATTTATTCCA
AGATCTGATGACGATTCATGTCGGTTGCACGTGTTTTGTCTTGAGCATGCTGCGGAAGTGGAACAG
CAACTTCGTCCTTTTGGGGGGATTAACTTAATGTTACTGTGCCATCCAGAGTACCCCAGGATAGAG
GCTGAAGCAAAGATAGTTGCCGAAGAGCTCGTCATCAATCACGAATGGAATGATACTGAATTCAGG
AATGTGACCCGAGAGGATGAGGAAACGATTCAGGCAGCGTTGGATAATGTTGAAGCTAAGGGTGGG
AACAGTGATTGGACCGTAAAATTGGGTGTTAACCTTTCTTACAGCGCTATTCTCAGTCGCTCTCCT
CTGTACAGTAAGCAGATGCCGTATAACTCCATCATATACAAGGCGTTCGGTCGCAGCTCTCCAGTA
GCGAGCTCACCCTCGAAACCCAAAGTCTCTGGTAAAAGATCGTCCAGACAGAGGAAATATGTTGTT
GGAAAATGGTGTGGTAAGGTTTGGATGTCACATCAGGTGCATCCCTTTTTGCTGGAGCAGGACTTA
GAGGGGGAAGAATCTGAAAGAAGTTGTCATCTTCGAGTTGCTATGGATGAGGATGCCACTGGAAAG
AGATCGTTTCCTAATAATGTTTCCAGGGATTCGACAACAATGTTTGGAAGAAAGTATTGTAGGAAG
AGAAAGATAAGAGCAAAGGCGGTGCCACGCAAGAAGCTTACTTCTTTTAAGAGGGAAGATGGAGTT
TCTGATGACACATCAGAAGATCATTCTTATAAGCAGCAATGGAGGGCTTCCGGGAATGAGGAAGAG
TCTTATTTTGAGACAGGGAACACAGCTTCTGGTGATTCATCAAATCAAATGTCTGATCCGCACAAG
GGAATTATCAGACATAAAGGTTATAAAGAATTTGAGTCAGATGATGAGGTTTCAGACCGTTCACTT

FIGURA 10 CONT.

GGGGAAGAGTATACTGTACGGGCATGTGCAGCTTCAGAGAGCTCAATGGAGAATGGCTCTCAGCAT
TCAATGTATGACCATGATGATGATGATGATGATATCGACAGGCAGCCTAGGGGGATTCCAAGGAGC
CAACAGACAAGAGTTTTTAGGAATCCAGTTTCATATGAGTCAGAAGATAATGGCGTTTATCAGCAA
AGCGGAAGAATATCCATAAGTAATAGGCAAGCTAATCGAATGGTTGGTGAATATGATTCAGCAGAG
AATTCTTTGGAGGAACGAGGCTTTTGCAGTACAGGGAAAAGGCAAACCAGGTCAACAGCCAAACGA
ATAGCAAAAACCAAGACAGTTCAGAGTTCGAGAGACACAAAAGGTCGCTTTTTGCAAGAATTTGCA
TCTGGAAAGAAGAATGAAGAATTGGATTCATACATGGAGGGACCTAGCACACGGCTTAGGGTGAGA
CATCAGAAGCCGTCGAGAGGGTCTTTAGAAACAAAACCAAAGAAGATTGGTAAGAAGAGAAGTGGT
AATGCTTCCTTCTCCAGAGTTGCAACTGAAAAAGATGTGGAGGAAAAAGAAGAAGAAGAAGAAGAA
GAAGAGAATGAGGAAGAGGAATGTGCAGCATACCAATGTAACATGGAGGGTTGCACGATGAGTTTC
AGTTCGGAAAAACAGTTGATGTTACACAAAAGAAACATATGCCCAATTAAAGGCTGTGGTAAAAAC
TTCTTCTCACACAAGTATTTGGTTCAACACCAGCGTGTTCACTCAGACGACCGTCCTCTGAAATGT
CCATGGAAGGGATGTAAGATGACGTTCAAGTGGGCTTGGTCTAGAACCGAGCACATAAGGGTTCAC
ACAGGCGCTAGGCCTTATGTTTGCGCTGAACCGGATTGTGGTCAAACATTCAGGTTTGTCTCTGAC
TTCAGCCGGCATAAAAGGAAGACCGGTCATTCGGTTAAGAAGACCAACAAAAGGTGA

SEQ ID NO: 120, Proteína - *Arabidopsis thaliana*

MAVSEQSQDVFPWLKSLPVAPEFRPTLAEFQDPIAYILKIEEEASRYGICKILPPLPPPSKKTSIS
NLNRSLAARAAARVRDGGFGACDYDGGPTFATRQQQIGFCPRKQRPVQRPVWQSGEEYSFGEFEFK
AKNFEKNYLKKCGKKSQLSALEIETLYWRATVDKPFSVEYANDMPGSAFIPLSLAAARRRESGGEG
GTVGETAWNMRAMSRAEGSLLKFMKEEIPGVTSPMVYVAMMFSWFAWHVEDHDLHSLNYLHMGAGK
TWYGVPKDAALAFEEVVRVHGYGEELNPLVTFSTLGEKTTVMSPEVFVKAGIPCCRLVQNPGEFVV
TFPGAYHSGFSHGFNFGEASNIATPEWLRMAKDAAIRRAAINYPPMVSHLQLLYDFVLALGSRVPT
SINPKPRSSRLKDKARSEGERLTKKLFVQNIIHNNELLSSLGKGSPVALLPQSSSDISVCSDLRIG
SHLITNQENPIQLKCEDLSSDSVVVDLSNGLKDTVSVKEKFTSLCERSRNHLASTEKDTQETLSDA
ERRKNDAAVALSDQRLFSCVTCGVLSFDCVAIVQPKEAAARYLMSADCSFFNDWTAASGSANLGQA
ARSLHPQSKEKHDVNYFYNVPVQTMDHSVKTGDQKTSTTSPTIAHKDNDVLGMLASAYGDSSDSEE
EDQKGLVTPSSKGETKTYDQEGSDGHEEARDGRTSDFNCQRLTSEQNGLSKGGKSSLLEIALPFIP
RSDDDSCRLHVFCLEHAAEVEQQLRPFGGINLMLLCHPEYPRIEAEAKIVAEELVINHEWNDTEFR
NVTREDEETIQAALDNVEAKGGNSDWTVKLGVNLSYSAILSRSPLYSKQMPYNSIIYKAFGRSSPV
ASSPSKPKVSGKRSSRQRKYVVGKWCGKVWMSHQVHPFLLEQDLEGEESERSCHLRVAMDEDATGK
RSFPNNVSRDSTTMFGRKYCRKRKIRAKAVPRKKLTSFKREDGVSDDTSEDHSYKQQWRASGNEEE
SYFETGNTASGDSSNQMSDPHKGIIRHKGYKEFESDDEVSDRSLGEEYTVRACAASESSMENGSQH
SMYDHDDDDDDIDRQPRGIPRSQQTRVFRNPVSYESEDNGVYQQSGRISISNRQANRMVGEYDSAE
NSLEERGFCSTGKRQTRSTAKRIAKTKTVQSSRDTKGRFLQEFASGKKNEELDSYMEGPSTRLRVR
HQKPSRGSLETKPKKIGKKRSGNASFSRVATEKDVEEKEEEEEEENEEEECAAYQCNMEGCTMSF
SSEKQLMLHKRNICPIKGCGKNFFSHKYLVQHQRVHSDDRPLKCPWKGCKMTFKWAWSRTEHIRVH
TGARPYVCAEPDCGQTFRFVSDFSRHKRKTGHSVKKTNKR

SEQ ID NO: 121, DNA - *Arabidopsis thaliana*

ATGGAGAAAATGAGAGGGAAGCGAATCAGACCAAGAGATTCCGGCGAGTTAGTCGAAGATGGGAGA
TCTGAAAGCGAGCGTAAGACGAGGAAGAAAGAGAATGATGTTGTGAGCAAAGGAAGAATCGGAAGA
GGAAGAGGAAGAGGAGAGGTTTCGAAGCGATCTATTGAAATAGACATTTCTAATCCCGAGAAGGAT
ATTAAGCCTGACGGAAGCAGAAAGTGTTTGGGTTCGACGTGTCATCATTGTAAGATCTTGACGAGT
GAGAGTGATCTTATCTTCTGTTCAAAGTGTAACAAGAAGTGTTATTGCTTTGACTGCATCAAGAGA
TCGTATTCAGAAAGAACACATGAGGAAGTTAGAGCTGCTTGTCCATTCTGTATGATGACTTGCATT
TGCAGAGCTTGTTTGCGGTTGCCTTTGGTTATCAAGCCTCCAAGTGAAAAAGATACAGATGTCAAG
CTTAAGCAGCTGCAGTATCTGTTAGTCAAAGTTCTCCCAGTTCTTAAGGATATTTATACGGAGCAG

FIGURA 10 CONT.

AACCGTGAACTAGAGATTGAATCAACAATTAGAGGACACCCTGTGACAGAAGCTAATATCAAGAGG
TGCAAACTTGACCCAAGCGAACGCATATACTGTGATCTCTGCCGCACATCCATTGCCAATTTTCAC
AGAAGCTGTCCGAACAAAAATTGTTCAGTCGATATATGTCTTTCTTGCTGCAAGGAACTAAGTGAG
GGCTTCCATCAAGAGAGAGATGGGAAGAAAAATGCAGAAGGGAAAGGATATGAATGTCGAATCCCA
GCAGGTCAAGGAAAAGACTCGGACGCATATGTTCCACTGCATTTCTCGACTTGGAAGCTTAACTCA
GACAGTAGCATTCCATGCCCTCCAAAAGAGTGTGGTGGTTGTGGTACTTCAACACTGGAGTTGAGA
CGCTTATGGAAAAGAGATTGGGTTGAAAAATTAATAACGAATGCCGAGAAATGTACTCTGAATTTT
AGGCCAACAGATGTGGATATTGTTCATGAATGTTCCTCATGCTCTACCAATTCTGATTCCATTAGA
AGGCAGGCAGCTTTTAGGAAGAATGCTCATGATAATTTCTTATACTCCCCAAATGCTGTTGATCTG
GCTGAAGACGATATTGCTCATTTTCAGTTTCATTGGATGAAGGCAGAACCTGTGATAGTCAGAAAT
GTTCTCGAGAAGACATCTGGACTAAGTTGGGAACCAATGGTTATGTGGAGGGCTTGTAGGGAAATG
GATCCAAAGCGTAAAGGTACAGAAGAAGAGACAACAAAAGTGAAAGCTTTGGATTGCTTGGACTGG
TGTGAGGTTGAAATAAATCTGCATCAGTTTTTTGAGGGCTACTTAGAAGGCCGCATGCATAAAAAC
GGTTGGCCAGAAATGTTGAAATTGAAGGACTGGCCTCCCTCAGATTTATTTGAAAAGCGCCTTCCA
AGGCATAATGCTGAGTTTATTGCCGCATTGCCTTTCTTTGATTACACTGACCCAAAGTCTGGTATC
CTAAATCTCGCAACAAGATTTCCGGAAGGATCCTTGAAGCCAGATCTTGGACCCAAGACATACATT
GCGTACGGTTTCCATGAAGAGCTTAATAGAGGAGATTCGGTGACAAAACTTCATTGTGACATTTCT
GACGCGGTCAATGTGTTGACACACACAGCTAAAGTGGAAATCCCTCCCGTCAAATACCAAAACATA
AAAGTACATCAGAAAAAATATGCAGAAGCCATGTTGCAGAAACAACAATACAGTGGTCAAGTGAAA
GAAGCCAGCGAACTTGAAAACAAGTCAATGAAAGAAGTGGATGAAAGTAAAAAGGATTTAAAAGAC
AAGGCAGCTAACGAAGAACAATCAAATAACAGTTCAAGACCGTCGGGCTCAGGAGAAGCTGAGAAA
GTAATTATTTCAAAAGAGGATAATCCCACACAACCTGCTGTGAGTACAAGTGTGGAATCCATACAA
GAACAGAAGCTGGATGCTCCAAAAGAAACTGATGGCAATACCAATGAGAGGTCAAAGGCTGTGCAT
GGTGGTGCTGTCTGGGATATCTTTCGAAGAGAGGATGTTCCAAAACTTATTCAGTTTTTGAAAAGG
CACGAGCATGAGTTTCGTCACTTCAATAACGAACCCCTGGAATCCGTTATTCACCCAATTCATGAC
CAGACTATGTTCTTGAGTGATAGCCAGAAGAAACAGCTGAAAGAAGAATTTGATATAGAACCATGG
ACCTTTGAGCAACACCTCGGCGAAGCAGTCTTCATACCTGCAGGTTGTCCTCACCAAGTGAGGAAT
AGACAGTCTTGCATAAAGGTGGCACTAGACTTTGTGGCTCCTGAAAGCGTCGAAGAATGCCTTAGG
CTAACACAGGAGTTCCGGAGATTGCCTAAAGACCACAGTTCTAGCGAGGATAAATTGGAGCTTAAG
AAAATTGCACTGTATGCTGCGAGTTCAGCCATAAGAGAGGTAAAAGGGCTAATGCAAAGCTCCCGA
CGCAGTGATACCTAA

SEQ ID NO: 122, Proteína - *Arabidopsis thaliana*

MEKMRGKRIRPRDSGELVEDGRSESERKTRKKENDVVSKGRIGRGRGRGEVSKRSIEIDISNPEKD
IKPDGSRKCLGSTCHHCKILTSESDLIFCSKCNKKCYCFDCIKRSYSERTHEEVRAACPFCMMTCI
CRACLRLPLVIKPPSEKDTDVKLKQLQYLLVKVLPVLKDIYTEQNRELEIESTIRGHPVTEANIKR
CKLDPSERIYCDLCRTSIANFHRSCPNKNCSVDICLSCCKELSEGFHQERDGKKNAEGKGYECRIP
AGQGKDSDAYVPLHFSTWKLNSDSSIPCPPKECGGCGTSTLELRRLWKRDWVEKLITNAEKCTLNF
RPTDVDIVHECSSCSTNSDSIRRQAAFRKNAHDNFLYSPNAVDLAEDDIAHFQFHWMKAEPVIVRN
VLEKTSGLSWEPMVMWRACREMDPKRKGTEEETTKVKALDCLDWCEVEINLHQFFEGYLEGRMHKN
GWPEMLKLKDWPPSDLFEKRLPRHNAEFIAALPFFDYTDPKSGILNLATRFPEGSLKPDLGPKTYI
AYGFHEELNRGDSVTKLHCDISDAVNVLTHTAKVEIPPVKYQNIKVHQKKYAEAMLQKQQYSGQVK
EASELENKSMKEVDESKKDLKDKAANEEQSNNSSRPSGSGEAEKVIISKEDNPTQPAVSTSVESIQ
EQKLDAPKETDGNTNERSKAVHGGAVWDIFRREDVPKLIQFLKRHEHEFRHFNNEPLESVIHPIHD
QTMFLSDSQKKQLKEEFDIEPWTFEQHLGEAVFIPAGCPHQVRNRQSCIKVALDFVAPESVEECLR
LTQEFRRLPKDHSSSEDKLELKKIALYAASSAIREVKGLMQSSRRSDT

FIGURA 10 CONT.

SEQ ID NO: 123, DNA - *Arabidopsis thaliana*

ATGGATCAGCTTGCATCTCTAGCAGAGTCTGTGGCTATGGAGGAAGATTCTGAGAAACAATCGATT
AAAGGAGAGAGTAGTCTTGAACCTGACTCTACTCCTTCTAGTCCTAAAATAACTGCTAGGTGGAAT
CCATCAGAAGCTTGCAGGCCTTTGGTTGATGACGCTCCCATCTTTTATCCCACTAATGAGGATTTT
GATGATCCACTTGGTTACATAGAGAAGTTGCGCTCTAAAGCAGAATCATATGGCATATGTCGAATT
GTGCCGCCTGTTGCATGGAGGCCCCCTTGCCCTCTGAAGGAGAAAAAAATATGGGAGAATTCTAAA
TTTCCCACCCGGATTCAGTTCATTGACTTGCTTCAAAACCGAGAACCGATTAAAAAATCTACTAAA
ACCAAGAAGAGAAAACGGAGAAGAATCTCTAAAATTGGATACACAAGAAGAAAAAGAGATTCAGGC
TGTGATACGGCTTCCTCTGGCTCTAGTGATAGTGAAGGAAAATTTGGTTTTCAGACTGGGCCAGAT
TTTACCCTGGAAGAATTTCAGAAGTATGATGAATACTTTAAGGAATGCTATTTTCAGTCAGAGGAT
CATCCTGGTTCCAAAGCGTCTGAAAATAAGAAATTTAAACCTAAGGTTAAGGACCTTGAAGGTGAA
TATTGGCGGATAGTGGAGCAGGCAACTGATGAAGTAGAGGTGTACTATGGAGCTGACTTGGAAACA
AAGAAATTTGGAAGTGGTTTCCCAAAGTATAAACCGGGGTATCCTATAAGTGAAGCAGATCAATAC
TCTCAATGTGGTTGGAACCTAAATAATTTGTCCCGTCTGCCCGGATCAGTTCTAGCTTTTGAGAGC
TGCGATATCTCAGGAGTCATTGTGCCATGGCTTTATGTTGGAATGTGCTTTTCAACTTTTTGTTGG
CATGTTGAAGATCATCACCTTTATTCCATGAACTACTTACACACAGGTGATCCAAAAGTTTGGTAT
GGGATCCCTGGAAACCATGCTGAGTCTTTTGAAAATGTAATGAAAAAGCGTCTTCCAGATTTGTTT
GAAGAACAGCCTGACTTGCTACATCAACTGGTCACTCAGTTATCTCCAAGAATCTTAAAAGAGGAG
GGAGTACCTGTCTATCGAGCTGTCCAGCGCAGTGGAGAATTCATTCTAACCTTCCCTAAGGCCTAT
CATTCTGGGTTTAATTGTGGTTTCAACTGTGCGGAAGCAGTAAATGTTGCACCAGTTGATTGGCTG
GTTCATGGACAGAATGCAGTGGAAGGCTATAGCAAGCAGCGGCGAAAGAGTTCATTGTCACATGAC
AAGCTGCTTCTTGGAGCTGCTATGGAAGCTACTTATTGCCTCTGGGAGCTTTCACTTTCAAAGAAG
AAGACTCCTGTGATTGCGAGATGGAAAAGGGTTTGTAGTGAGGATGGATTGCTTACGAAGGCAGTC
AAGAAGCGTGTGCAGATGGAAGAAGAAAGACTAAATCACCTTCAAGATGGTTTTAGCTTGCGAAAG
ATGGAGGGTGATTTCGACAATAAGAGAGAACGGGAGTGCTTCTTGTGCTTCTATGATTTGCATATG
TCTGCTTCAAGCTGCAAGTGTTCCCCCAACCGGTTTGCATGTCTTATCCACGCTAAGGATCTGTGT
TCATGTGAAAGTAAGGATAGATATATCCTAATTCGCCACACCTTGGATGAGTTATGGGCACTGGTC
AGAGCTCTAGAAGGGGATCTTGATGCCATAGACCTATGGGCAAGTAAATGTCGTGACCAGTATCCT
TCACAGCATCCAAGAGCAAGAGAATATGCTTACCTCAAGTCTGCCCCATGTATAAAGAGCCGTGGT
TCATCAAAAGTACAGCAGCGAGAACAAAACAATCTTCAGTTAGTGTCTGAACGCTTACAAAGTGAT
CTAACCTCAAACAAGGAAGTTCAGTTGAAACAAGATGGTGATTCAGATGTAAATCGCCATGGTCAT
GAAAGCGAAAGAAATCATGTACATGGGATCACTGATAAGTCAGCTGTCACGGATGTAAAACTTGGT
GTGGGAGGTAAGTTTGATGAGAAGAAGATTTCGGTTGAATCTCAAAACCCACATTCAGTTTCAGAT
GTAGGGTGTAGCGAGCTGGCGAAGAAAGTGGATGGTTGTTTAGGAGGGAAGGACCAAAATGCTGCA
ACCAATAGGTTAAGTCTCTCTGTTGAGCTTTTGAGTTCTGGATCTCTCGTTGTTAAAAAGCTGTGG
TGCAGTAAACAGGCCATATACCCAAAAGGGTTCAAGAGCCGTGTTAAGTTCTTAAGTGTGCTTGAC
CCAACAAACCTAACCAACTACATCTCAGAGGTTCTGGATGCAGGGCTTCTTGGTCCATTGTTCAGG
GTCTCAGTAGAAGATTACCCCACTGAGAATTTCTCGAATGTATCTGCTGAAAAGTGCTGGCAAATG
GTGACACAAAGGCTCAAACTCGAAATCATCAAGAAATGTGATCAACCAGTTAGCTCTTTGACCTCT
TTACAACCTTTGGAGAGTATAAATGGGCTTGAAATGTTTGGATTTCTCTCCCCGCACGTAATTAAG
GTGGTTGAGGCTCTTGATCCAAAGCATCAATTGGAGGAGTATTGGAACCAAAAAGCAGTGAAACTG
TTTGGTGCTGAACCAATAAAGGAAGGAGAAAAGGATGATACAGAGAAAGGAGGGGCTTCAGATCCC
TCTTTGGACCGGGACACAAGGCTTCTGCGTGGACTGTTGAAGAAGGCGACACCTGAAGAATTAGTA
ATGATGCATGGACTTCTGTGTGGTGAGACTCGCAACACCGAGCTCAAGGAAGAGCTCTCTACTTTG
GTTGATAAGATGGAGATAAGTCCTTAA

FIGURA 10 CONT.

SEQ ID NO: 124, Proteína - *Arabidopsis thaliana*

MDQLASLAESVAMEEDSEKQSIKGESSLEPDSTPSSPKITARWNPSEACRPLVDDAPIFYPTNEDF
DDPLGYIEKLRSKAESYGICRIVPPVAWRPPCPLKEKKIWENSKFPTRIQFIDLLQNREPIKKSTK
TKKRKRRRISKIGYTRRKRDSGCDTASSGSSDSEGKFGFQTGPDFTLEEFQKYDEYFKECYFQSED
HPGSKASENKKFKPKVKDLEGEYWRIVEQATDEVEVYYGADLETKKFGSGFPKYKPGYPISEADQY
SQCGWNLNNLSRLPGSVLAFESCDISGVIVPWLYVGMCFSTFCWHVEDHHLYSMNYLHTGDPKVWY
GIPGNHAESFENVMKKRLPDLFEEQPDLLHQLVTQLSPRILKEEGVPVYRAVQRSGEFILTFPKAY
HSGFNCGFNCAEAVNVAPVDWLVHGQNAVEGYSKQRRKSSLSHDKLLLGAAMEATYCLWELSLSKK
KTPVIARWKRVCSEDGLLTKAVKKRVQMEEERLNHLQDGFSLRKMEGDFDNKRERECFLCFYDLHM
SASSCKCSPNRFACLIHAKDLCSCESKDRYILIRHTLDELWALVRALEGDLDAIDLWASKCRDQYP
SQHPRAREYAYLKSAPCIKSRGSSKVQQREQNNLQLVSERLQSDLTSNKEVQLKQDGDSDVNRHGH
ESERNHVHGITDKSAVTDVKLGVGGKFDEKKISVESQNPHSVSDVGCSELAKKVDGCLGGKDQNAA
TNRLSLSVELLSSGSLVVKKLWCSKQAIYPKGFKSRVKFLSVLDPTNLTNYISEVLDAGLLGPLFR
VSVEDYPTENFSNVSAEKCWQMVTQRLKLEIIKKCDQPVSSLTSLQPLESINGLEMFGFLSPHVIK
VVEALDPKHQLEEYWNQKAVKLFGAEPIKEGEKDDTEKGGASDPSLDRDTRLLRGLLKKATPEELV
MMHGLLCGETRNTELKEELSTLVDKMEISP

SEQ ID NO: 125, DNA - *Arabidopsis thaliana*

ATGATACTGCTTCATGGTCAGGATTTTGATGATCCACTTGGTTACATAGAGAAGTTGCGCTCTAAA
GCAGAATCATATGGCATATGTCGAATTGTGCCGCCTGTTGCATGGAGGCCCCCTTGCCCTCTGAAG
GAGAAAAAAATATGGGAGAATTCTAAATTTCCCACCCGGATTCAGTTCATTGACTTGCTTCAAAAC
CGAGAACCGATTAAAAAATCTACTAAAACCAAGAAGAGAAAACGGAGAAGAATCTCTAAAATTGGA
TACACAAGAAGAAAAAGAGATTCAGGCTGTGATACGGCTTCCTCTGGCTCTAGTGATAGTGAAGGA
AAATTTGGTTTTCAGACTGGGCCAGATTTTACCCTGGAAGAATTTCAGAAGTATGATGAATACTTT
AAGGAATGCTATTTTCAGTCAGAGGATCATCCTGGTTCCAAAGCGTCTGAAAATAAGAAATTTAAA
CCTAAGGTTAAGGACCTTGAAGGTGAATATTGGCGGATAGTGGAGCAGGCAACTGATGAAGTAGAG
GTGTACTATGGAGCTGACTTGGAAACAAAGAAATTTGGAAGTGGTTTCCCAAAGTATAAACCGGGG
TATCCTATAAGTGAAGCAGATCAATACTCTCAATGTGGTTGGAACCTAAATAATTTGTCCCGTCTG
CCCGGATCAGTTCTAGCTTTTGAGAGCTGCGATATCTCAGGAGTCATTGTGCCATGGCTTTATGTT
GGAATGTGCTTTTCAACTTTTTGTTGGCATGTTGAAGATCATCACCTTTATTCCATGAACTACTTA
CACACAGGTGATCCAAAAGTTTGGTATGGGATCCCTGGAAACCATGCTGAGTCTTTTGAAAATGTA
ATGAAAAAGCGTCTTCCAGATTTGTTTGAAGAACAGCCTGACTTGCTACATCAACTGGTCACTCAG
TTATCTCCAAGAATCTTAAAAGAGGAGGGAGTACCTGTCTATCGAGCTGTCCAGCGCAGTGGAGAA
TTCATTCTAACCTTCCCTAAGGCCTATCATTCTGGGTTTAATTGTGGTTTCAACTGTGCGGAAGCA
GTAAATGTTGCACCAGTTGATTGGCTGGTTCATGGACAGAATGCAGTGGAAGGCTATAGCAAGCAG
CGGCGAAAGAGTTCATTGTCACATGACAAGCTGCTTCTTGGAGCTGCTATGGAAGCTACTTATTGC
CTCTGGGAGCTTTCACTTTCAAAGAAGAAGACTCCTGTGATTGCGAGATGGAAAAGGGTTTGTAGT
GAGGATGGATTGCTTACGAAGGCAGTCAAGAAGCGTGTGCAGATGGAAGAAGAAAGACTAAATCAC
CTTCAAGATGGTTTTAGCTTGCGAAAGATGGAGGGTGATTTCGACAATAAGAGAGAACGGGAGTGC
TTCTTGTGCTTCTATGATTTGCATATGTCTGCTTCAAGCTGCAAGTGTTCCCCCAACCGGTTTGCA
TGTCTTATCCACGCTAAGGATCTGTGTTCATGTGAAAGTAAGGATAGATATATCCTAATTCGCCAC
ACCTTGGATGAGTTATGGGCACTGGTCAGAGCTCTAGAAGGGGATCTTGATGCCATAGACCTATGG
GCAAGTAAATGTCGTGACCAGTATCCTTCACAGCATCCAAGAGCAAGAGAATATGCTTACCTCAAG
TCTGCCCCATGTATAAAGAGCCGTGGTTCATCAAAAGTACAGCAGCGAGAACAAAACAATCTTCAG
TTAGTGTCTGAACGCTTACAAAGTGATCTAACCTCAAACAAGGAAGTTCAGTTGAAACAAGATGGT
GATTCAGATGTAAATCGCCATGGTCATGAAAGCGAAAGAAATCATGTACATGGGATCACTGATAAG
TCAGCTGTCACGGATGTAAAACTTGGTGTGGGAGGTAAGTTTGATGAGAAGAAGATTTCGGTTGAA
TCTCAAAACCCACATTCAGTTTCAGATGTAGGGTGTAGCGAGCTGGCGAAGAAAGTGGATGGTTGT

FIGURA 10 CONT.

TTAGGAGGGAAGGACCAAAATGCTGCAACCAATAGGTTAAGTCTCTCTGTTGAGCTTTTGAGTTCT
GGATCTCTCGTTGTTAAAAAGCTGTGGTGCAGTAAACAGGCCATATACCCAAAAGGGTTCAAGAGC
CGTGTTAAGTTCTTAAGTGTGCTTGACCCAACAAACCTAACCAACTACATCTCAGAGGTTCTGGAT
GCAGGGCTTCTTGGTCCATTGTTCAGGGTCTCAGTAGAAGATTACCCCACTGAGAATTTCTCGAAT
GTATCTGCTGAAAAGTGCTGGCAAATGGTGACACAAAGGCTCAAACTCGAAATCATCAAGAAATGT
GATCAACCAGTTAGCTCTTTGACCTCTTTACAACCTTTGGAGAGTATAAATGGGCTTGAAATGTTT
GGATTTCTCTCCCCGCACGTAATTAAGGTGGTTGAGGCTCTTGATCCAAAGCATCAATTGGAGGAG
TATTGGAACCAAAAAGCAGTGAAACTGTTTGGTGCTGAACCAATAAAGGAAGGAGAAAAGGATGAT
ACAGAGAAAGGAGGGGCTTCAGATCCCTCTTTGGACCGGGACACAAGGCTTCTGCGTGGACTGTTG
AAGAAGGCGACACCTGAAGAATTAGTAATGATGCATGGACTTCTGTGTGGTGAGACTCGCAACACC
GAGCTCAAGGAAGAGCTCTCTACTTTGGTTGATAAGATGGAGATAAGTCCTTAA

SEQ ID NO: 126, Proteína - *Arabidopsis thaliana*

MILLHGQDFDDPLGYIEKLRSKAESYGICRIVPPVAWRPPCPLKEKKIWENSKFPTRIQFIDLLQN
REPIKKSTKTKKRKRRRISKIGYTRRKRDSGCDTASSGSSDSEGKFGFQTGPDFTLEEFQKYDEYF
KECYFQSEDHPGSKASENKKFKPKVKDLEGEYWRIVEQATDEVEVYYGADLETKKFGSGFPKYKPG
YPISEADQYSQCGWNLNNLSRLPGSVLAFESCDISGVIVPWLYVGMCFSTFCWHVEDHHLYSMNYL
HTGDPKVWYGIPGNHAESFENVMKKRLPDLFEEQPDLLHQLVTQLSPRILKEEGVPVYRAVQRSGE
FILTFPKAYHSGFNCGFNCAEAVNVAPVDWLVHGQNAVEGYSKQRRKSSLSHDKLLLGAAMEATYC
LWELSLSKKKTPVIARWKRVCSEDGLLTKAVKKRVQMEEERLNHLQDGFSLRKMEGDFDNKREREC
FLCFYDLHMSASSCKCSPNRFACLIHAKDLCSCESKDRYILIRHTLDELWALVRALEGDLDAIDLW
ASKCRDQYPSQHPRAREYAYLKSAPCIKSRGSSKVQQREQNNLQLVSERLQSDLTSNKEVQLKQDG
DSDVNRHGHESERNHVHGITDKSAVTDVKLGVGGKFDEKKISVESQNPHSVSDVGCSELAKKVDGC
LGGKDQNAATNRLSLSVELLSSGSLVVKKLWCSKQAIYPKGFKSRVKFLSVLDPTNLTNYISEVLD
AGLLGPLFRVSVEDYPTENFSNVSAEKCWQMVTQRLKLEIIKKCDQPVSSLTSLQPLESINGLEMF
GFLSPHVIKVVEALDPKHQLEEYWNQKAVKLFGAEPIKEGEKDDTEKGGASDPSLDRDTRLLRGLL
KKATPEELVMMHGLLCGETRNTELKEELSTLVDKMEISP

SEQ ID NO: 127, DNA - *Arabidopsis thaliana*

ATGGGTAATGTTGAAATTCCGAATTGGCTTAAAGCTTTGCCATTGGCACCTGTGTTTAGACCTACA
GACACTGAATTTGCAGATCCCATTGCTTACATATCGAAAATCGAAAAGGAAGCTAGTGCCTTTGGG
ATTTGCAAGATCATTCCTCCTTTACCCAAGCCATCGAAAAAGTATGTTTTCTACAACTTGAACAAG
TCTCTTTTGAAGTGTCCCGAATTGGTTTCAGATGTTGACATTTCGAAGGTGTGTAAAGAGGATAGA
GCTGTATTCACCACTAGGCAGCAAGAGTTAGGCCAAACTGTGAAAAAAAACAAAGGAGAGAAGGGT
AAGAGTAATTCTCAAAGGAGTGGTGTTAAGCAGGTTTGGCAAAGTGGAGGCGTTTATACATTGGAT
CAGTTTGAAGCTAAGTCGAAAGCTTTCTACAAAACCCAATTAGGAACGGTTAAGGAACTGGCACCA
GTTGTCATTGAGGCATTGTTCTGGAAAGCGGCTTTAGAGAAGCCTATATATATAGAGTATGCTAAT
GACGTGCCTGGCTCTGCTTTCGGTGAGCCAGAGGATCATTTCAGGCATTTCCGGCAGAGGAAGAGG
AGGGGCAGGGGATTTTATCAGAGGAAGACAGAAAATAATGATCCATCAGGTAAGAACGGTGAAAAG
TCATCGCCTGAGGTAGAAAAGGCACCCCTTGCTTCTACAAGTTTATCATCTCAGGATTCGTCCAAG
CAGAAGAACATGGATATTGTTGATGAAATGGAAGGTACTGCAGGCTGGAAGCTCTCCAACAGTTCG
TGGAACCTTCAGATGATTGCACGTTCACCTGGTTCTGTTACACGTTTCATGCCAGATGACATTCCT
GGTGTTACATCTCCCATGGTTTATATCGGTATGTTGTTCAGTTGGTTTGCCTGGCACGTTGAGGAT
CATGAGCTTCACAGTATGAATTATCTTCACACTGGCTCACCAAAGACTTGGTACGCTGTCCCTTGT
GATTATGCGCTTGACTTTGAAGAGGTTATCCGCAAAAATTCTTATGGCAGGAACATTGACCAACTG
GCTGCTCTCACCCAACTAGGCGAAAAGACAACTCTTGTATCACCTGAGATGATAGTTGCATCTGGC
ATTCCCTGCTGTAGGTTGGTACAAAATCCTGGTGAATTTGTTGTGACTTTTCCGAGGTCTTATCAT
GTAGGATTCAGCCACGGTTTTAACTGTGGGGAAGCCGCTAATTTTGGAACCCCACAGTGGCTCAAC

FIGURA 10 CONT.

GTAGCAAAGGAAGCTGCTGTGCGTCGGGCAGCCATGAATTATCTACCCATGCTGTCCCATCAGCAG
CTGCTATATCTCTTGACTATGTCCTTTGTTTCAAGAGTGCCACGATCATTGCTACCAGGTGGTCGT
AGCTCCCGACTAAGAGATCGGCAGAGAGAAGAAAGAGAGTTCCTTGTGAAAAGAGCTTTTGTAGAA
GATATATTGAACGAAAACAAGAATTTATCTGTTCTTCTTCGAGAACCAGGCTCTCGTTTGGTGATG
TGGGACCCTGATTTACTTCCGCGTCATAGCGCACTGGCTCTTGCAGCTGCTGGAGTTGCTGGTGCT
TCTGCTGTTTCACCTCCAGCAGTGGCAAAAAAAGAACTTGAGGAGGGTCATTCTGAATTGCAGAAT
AAAGAAAAGACTTCTCTTTTAGAGGAATTGAGCTTGTTCATGGAGAAGCTGAATGATGTATACTAC
GACGATGATGATGGTCTGCTTAACGATTTCCAAGTTGATACTGGAACCTTGCCATGTGTGGCTTGT
GGCGTTCTTGGCTTCCCCTTTATGTCTGTGGTACAGCCTTCTGAAAAAGCATTAAAAGATCTCTCA
GAGAGACAAGGAGAGACAGATGCTCAGGAAATTATGACACTGTCATCAGAAAAGTCTGACTGCGAA
TGGAAAACGTCTTCTAGATATATAAGACCTCGCATTTTCTGCCTCGAACACACTATTGAACTTCAG
AGACTGCTGCAATCAAGAGGTGGATTGAAATTCCTTGTTATTTGCCATAAAGACTTTCAAAAATTC
AAGGCACATGCGGCTATAGTGGCAGAGGAGGTCAAAGTCCCTTTCAGCTATGATGATGTCCTGTTA
GAGAGTGCATCTCAGGAAGAGTTGAGTCTAATTGATCTTGCAATTGAAGATGAAGAAAAATACGAA
CACAGTGTAGACTGGACCTCAGAACTTGGCATCAATTTACGGTACTGTGTGAAAGTGAGGAAAAAT
TCCCCTACTAAGAAAATTCAGCATGCACTGTCACTAGGTGGCTTGTTCTCCGATACAAGCCAGATG
CTAGATTTTACAACTATCAGATGGCTGCAGAGAAAATCACGCTCAAAAGCTAAACCCAGTTCTACC
TCAAGTTTCACACCTTGTGAACATCTTGAAGTAAAAGCAGATGGAAAATTAAGGGATAATTTGGAT
TCTCAGACCGGAAAAAAGGAAGAAAAGATCATCCAGTACTCGAGAAAGAAAAAGTTGAATCCAAAG
CCTTCAGCAGAACAAGTTCAGGAGCTAGCTACCCTAGCTAAATCAAAAGATTTTGATAAGACATGT
AAAAATTTTAGCAGTAGGTCACATCTGGATAGTGCAATTCGTTCTGAGATGAACAGTGAGATTGGA
GACTCTGGAAGGGTCATTGGGGTATCATTTTCCATAAATCCTTGTAGCTCATCTTTCACTGTGGGA
CATGGGCAAGAACATCCCGAGATCACTGTCAAGTTTGGTTCAGATTTAGATGGGAATGTCACAAAT
AGTTTGAGTATGGTGAATGGAGACTCTGCTGACCTAACCTTAACCTCCATATCCAGAGAGCAGCAC
CAAGGACACTCTATGACCAGCAATAATAATGGCTCCAACTCAGGTAGTCATGTTGTGGCAAGCCAG
ACCATACTTGTTTCTACAGGTGATAATCATGACGGACCAAGAAAATTGTCTGGTGATTATGTCTGC
AGTGATGTGTCTGTACGTGGTATTCAGGAAGCGGTTGAAATGAGTGACCAAGAGTTTGGAGAACCT
AGGTCTACTGTCACTAATATTGAGGATGAACAGCAATCACAGATTGTGAAACCAACCCAAAGAGAA
GCTGTATTTGGTGATCACGAACAGGTGGAGGGAGCAGAAGCTGTGTCTACCAGAGAAAACTTGTGC
TCTGAAATAATTCTGCACACTGAGCATTCTTCAGCTCATGTGGGTATGGAAATTCCTGACATAAAC
ACGGCCAGTGAGAACTTAGTTGTTGACATGACCCATGATGGTGAACCTTTGGAAAGTAGCGATATA
TTAAGTTCGAGCAATGGTGATGAAGCTTCTTCAAATGGATTGCAAGTTCTAAATGATGAGCTTAGC
ATGGAGAGTGAAGTTTCAAGCTCAGAAAACACCGAAGTTATTGAGGCTCCCAATTCTATGGGAGAA
GCAAAGAAGAAGCGGAAAATAGAATCAGAGTCTGAGACAAATGATAATCCAGAGAGTAGCATTGGT
TTCATAAGGAGTCCTTGTGAAGGGTTGAGGTCAAGGGGCAAGAGGAAAGCAACATGTGAAACTTCA
TTAAAGCACACTGAAACGAGTGATGAAGAGAAGAAACCCATTGCGAAAAGGCTGAAGAAAACACCA
AAGGCTTGCTCAGGGAGTCGTCAGCAAGAAGTCCCCACGACAACTCACCCCAACCGTTGTTACCTA
GAGGGATGCAAGATGACTTTTGAGAGTAAAGCGAAATTACAAACTCACAAGAGAAACCGTTGCACT
CACGAAGGGTGTGGAAAGAAATTCAGGGCTCACAAATATCTGGTGCTTCATCAACGTGTGCATAAG
GATGAAAGACCTTTCGAGTGCTCTTGGAAAGGATGTTCAATGACTTTCAAATGGCAATGGGCGAGG
ACAGAACATTTGCGTCTGCACACGGGAGAGCGACCATACATATGCAAGGTCGATGGATGTGGATTG
TCGTTTAGGTTTGTATCTGACTATAGTCGCCACAGACGTAAAACCATGCACTATGTCACATAG

SEQ ID NO: 128, Proteína - *Arabidopsis thaliana*

MGNVEIPNWLKALPLAPVFRPTDTEFADPIAYISKIEKEASAFGICKIIPPLPKPSKKYVFYNLNK
SLLKCPELVSDVDISKVCKEDRAVFTTRQQELGQTVKKNKGEKGKSNSQRSGVKQVWQSGGVYTLD
QFEAKSKAFYKTQLGTVKELAPVVIEALFWKAALEKPIYIEYANDVPGSAFGEPEDHFRHFRQRKR
RGRGFYQRKTENNDPSGKNGEKSSPEVEKAPLASTSLSSQDSSKQKNMDIVDEMEGTAGWKLSNSS

FIGURA 10 CONT.

WNLQMIARSPGSVTRFMPDDIPGVTSPMVYIGMLFSWFAWHVEDHELHSMNYLHTGSPKTWYAVPC
DYALDFEEVIRKNSYGRNIDQLAALTQLGEKTTLVSPEMIVASGIPCCRLVQNPGEFVVTFPRSYH
VGFSHGFNCGEAANFGTPQWLNVAKEAAVRRAAMNYLPMLSHQQLLYLLTMSFVSRVPRSLLPGGR
SSRLRDRQREEREFLVKRAFVEDILNENKNLSVLLREPGSRLVMWDPDLLPRHSALALAAAGVAGA
SAVSPPAVAKKELEEGHSELQNKEKTSLLEELSLFMEKLNDVYYDDDDGLLNDFQVDTGTLPCVAC
GVLGFPFMSVVQPSEKALKDLSERQGETDAQEIMTLSSEKSDCEWKTSSRYIRPRIFCLEHTIELQ
RLLQSRGGLKFLVICHKDFQKFKAHAAIVAEEVKVPFSYDDVLLESASQEELSLIDLAIEDEEKYE
HSVDWTSELGINLRYCVKVRKNSPTKKIQHALSLGGLFSDTSQMLDFTTIRWLQRKSRSKAKPSST
SSFTPCEHLEVKADGKLRDNLDSQTGKKEEKIIQYSRKKKLNPKPSAEQVQELATLAKSKDFDKTC
KNFSSRSHLDSAIRSEMNSEIGDSGRVIGVSFSINPCSSSFTVGHGQEHPEITVKFGSDLDGNVTN
SLSMVNGDSADLTLTSISREQHQGHSMTSNNNGSNSGSHVVASQTILVSTGDNHDGPRKLSGDYVC
SDVSVRGIQEAVEMSDQEFGEPRSTVTNIEDEQQSQIVKPTQREAVFGDHEQVEGAEAVSTRENLC
SEIILHTEHSSAHVGMEIPDINTASENLVVDMTHDGEPLESSDILSSSNGDEASSNGLQVLNDELS
MESEVSSSENTEVIEAPNSMGEAKKKRKIESESETNDNPESSIGFIRSPCEGLRSRGKRKATCETS
LKHTETSDEEKKPIAKRLKKTPKACSGSRQQEVPTTTHPNRCYLEGCKMTFESKAKLQTHKRNRCT
HEGCGKKFRAHKYLVLHQRVHKDERPFECSWKGCSMTFKWQWARTEHLRLHTGERPYICKVDGCGL
SFRFVSDYSRHRRKTMHYVT

SEQ ID NO: 129, DNA - *Arabidopsis thaliana*

ATGCCAAAGTGCAAGAATCTGTTGCTAACCTCGAAGCGACGTAAATCCAAATCCAAAAGACTAAAG
CTCCATCAGCATGAACCTGAGTCTCTGTTTCCAGAGAAGGAGGTAGAAGAGGAAGATGAAGATGAA
GGAGGCTTCAAGCTGAAAATTGCAGCTCCTTCTCAAGAGCATGGAGTTCAACCTCTGGGAAACCTA
TATTTCAATCCTGGTGCGGTCAATGTGCGAAACACTGGATTAGGTAATCTCCAGATACTCTCAGAT
GAGCTCGTTTTGGATATTCTAGGTCTTCTTGGCGCCAATCATTTGGGTGTTTTGGCTACTGTAACT
AAGTCATTCTATATATTTGCTAATCATGAGCCTCTCTGGAGAAATCTTGTGTTAGAGGAGTTAAAA
GGAGATTTTTTGTTTAATGGGTCGTGGAGATCTACCTATGTAGCTGCTTACCATCCGAAATTCAAG
TTTGCTGGAGATGGTGAATCGAATTTGAAGATTATAGATTTCTATTCTGATTATCTGTTTCAGAGC
TGGTTATGTGCGAATCTTGAAATGAAACCGAAATGGCTTAGAAGAGATAATATCACACGGGTGAGA
GGCATTTCTGTTGAAGACTTCATTACAAAGTTTGAGGAGCCGAATAAGCCAGTTTTGTTGGAGGGA
TGTTTGGATGGTTGGCCTGCTATTGAGAAATGGTCAAGAGATTATCTGACTAAGGTGGTTGGTGAT
GTCGAATTCGCGGTAGGTCCAGTGGAAATGAAGCTCGAGAAGTATTTTAGGTATTCTGATGGAGCT
AGGGAAGAAAGGCCTTTGTACTTGTTTGATCCAAAGTTTGCAGAGAAAGTTCCGGTTTTGGATTCG
GAGTATGATGTTCCTGTCTATTTCAGAGAGGACTTGTTTGGTGTTTTAGGAAACGAGCGACCTGAT
TATAGATGGATCATAATTGGTCCAGCCGGATCAGGATCGTCTTTCCATATTGATCCTAACTCAACA
TCAGCTTGGAATGCGGTGATCACAGGGTCGAAGAAATGGGTTTTGTTCCCACCTGATGTGGTTCCT
CCGGGTGTGCATCCAAGTCCTGATGGAGCAGAAGTGGCCTGTCCTGTTTCCATAATAGAATGGTTC
ATGAACTTTTATGATGATACTAAGGATTGGGAGAAGAAACCTATTGAGTGTATATGCAAAGCCGGG
GAGGTGATGTTTGTACCTAATGGATGGTGGCATCTGGTGATTAACCTGGAGGAATCGATTGCTATT
ACTCAGAATTATGCTAGCAGGAGTAATTTGTTGAATGTGTTAGAATTTCTGAAGAAACCAAATGCT
AAAGAACTTGTCTCCGGGACAACTGACAGAGAGAATTTGCATGACAAATTCAAGAAAGCCATTGAA
GAAGCTTACCCGGGGACTATCCAAGAGTTAGAAAAGAAAGCAGAAGAAGCAAAAAGAGCCGAAGAA
CAAAGAGTTTCTTTCTGGGACTCTGCCAAAACTGATACTTTCAAGTTTTCTTTCTAA

SEQ ID NO: 130, Proteína - *Arabidopsis thaliana*

MPKCKNLLLTSKRRKSKSKRLKLHQHEPESLFPEKEVEEEDEDEGGFKLKIAAPSQEHGVQPLGNL
YFNPGAVNVRNTGLGNLQILSDELVLDILGLLGANHLGVLATVTKSFYIFANHEPLWRNLVLEELK
GDFLFNGSWRSTYVAAYHPKFKFAGDGESNLKIIDFYSDYLFQSWLCANLEMKPKWLRRDNITRVR
GISVEDFITKFEEPNKPVLLEGCLDGWPAIEKWSRDYLTKVVGDVEFAVGPVEMKLEKYFRYSDGA

FIGURA 10 CONT.

REERPLYLFDPKFAEKVPVLDSEYDVPVYFREDLFGVLGNERPDYRWIIIGPAGSGSSFHIDPNST
SAWNAVITGSKKWVLFPPDVVPPGVHPSPDGAEVACPVSIIEWFMNFYDDTKDWEKKPIECICKAG
EVMFVPNGWWHLVINLEESIAITQNYASRSNLLNVLEFLKKPNAKELVSGTTDRENLHDKFKKAIE
EAYPGTIQELEKKAEEAKRAEEQRVSFWDSAKTDTFKFSF

SEQ ID NO: 131, DNA - *Arabidopsis thaliana*

ATGAAGGAAAAAGTGGAGAAGCTTGAGACAGATGATCTAAAATGGACCGAGAGGCTTCCAGAATGT
CCTGTCTACAGACCAACAAAGGAGGAGTTTGAGGATCCTCTGACTTATTTGCAGAAGATTTTCCCA
GAAGCTTCAAAATATGGTATTTGCAAGATTGTATCTCCTTTGACAGCAACAGTTCCTGCTGGCGCT
GTGTTGATGAAAGAGAAATCAAACTTTAAGTTCACTACCAGAGTACAACCCCTTCGACTTGCTGAG
TGGGATTCTGATGATAAAGTTACATTCTTCATGAGTGGGAGGACCTATACATTTCGCGACTATGAG
AAAATGGCGAACAAGGTATTTGCACGCAGATATTGCAGTGGTGGCTCTTTGCCCGACTCATTCTTG
GAGAAAGAATTCTGGAAGGAAATTGCGTGTGGCAAGACTGAAACAGTTGAGTATGCGTGTGATGTA
GATGGTAGCGCATTTTCATCTGCACCAGGTGACCCACTAGGAAGCAGCAAATGGAACTTGAATAAA
GTTTCAAGGCTTCCTAAGTCTACCCTGCGCCTTCTTGAAACATCCATTCCGGGAGTCACAGAACCC
ATGCTTTACATCGGAATGCTTTTCAGTATGTTTGCCTGGCATGTTGAGGACCATTATTTGTACAGC
ATTAATTACCAACATTGTGGAGCATCAAAAACTTGGTATGGGATTCCAGGTTCCGCAGCTCTAAAA
TTTGAAAAGGTGGTTAAAGAGTGTGTGTACAATGATGATATTCTATCAACTAATGGAGAAGATGGT
GCCTTTGACGTGCTTCTGGGGAAGACAACTATATTCCCCCCAAAGACTCTGTTGGATCATAATGTG
CCAGTGTATAAAGCTGTGCAAAAACCTGGAGAATTTGTTGTCACATTCCCCAGGGCTTATCATGCT
GGTTTCAGCCACGGTTTTAATTGTGGTGAGGCAGTGAACTTTGCGATGGGTGATTGGTTTCCATTT
GGAGCTATTGCCAGTTGCCGGTATGCTCATCTTAACCGAGTACCGTTGCTTCCTCATGAAGAATTG
ATTTGTAAAGAGGCAATGCTCTTGAACTCAAGTTCCAAATCCGAGAATCTGGATCTTACACCTACA
GAATTGTCAGGACAACGAAGCATCAAGACTGCATTTGTACACCTGATTCGTTTTCTTCATCTTGCT
CGATGGTCTCTAATGAAGTCAGGATTATGCACAGGCCTTGTTTCGAATACATATGGAACCATAGTC
TGCAGTTTGTGTAAACGGGATTGCTATTTGGCATTTATCAACTGCGAGTGTTACTCACACCCCGTT
TGTCTACGTCATGATGTTAAAAAGCTTGACCTTCCATGTGGGACAACGCATACTCTTTACTTGAGA
GACAACATTGAAGACATGGAAGCTGCTGCAATGAAGTTTGAGAAGGAAGATGGAGTTTCAGATTTG
ATTACAACTGATGAAGATTTATATAAGTATCCATCATCGATTACACTTCCAGCTGCGAAAGAAGAC
GGATATACACCGTATTCCACTATATACTTTGATTTCTATACTGAGGTAGAGATGACATCTCATGAC
CAGTTGCAATCGGGAAATCCCGTCATGAGTTATGAGGCAAACGCTTCATGTATCTCCTCAGTTGCT
GATGATTATGAATGCTCTGATTACCAGGTAAATAGACGAGCTAACTGTAGCAGCAGTAGTGATTCA
AAGCTCTCGGAAGAAGTAGCATGCAGCAGTAGCAAAAAAACTCGGTTCTTCCCTGTGGTTCAAGAT
GAACAACTAGTTGCGGATCAGGAAAGTGATGGATCTGATTCCGAGTGTTTTAGAGTCAAACGCCGT
TCTTCATTGAAGTTTGAGAACCGAACAGTTGTTCTAGACACAAGAGAGTCTGACCATCATCAGGAA
CTTAAGAGACTGAAGAAATCGCATCATCACGAAGGAAGATATAGTAGTAGCAGTAGTGTTAGTAGG
CAAGAAGAAGAAGAAGATGAGTTAGTAATTTCAAACAGGAAAGAAACTCAGCAACAAAGTGACGTG
AAGATGCAGAAGAAGAGGATCGAGAATCATTTTGGTGGCTTTAAACGTCTGAAAGTGAAAGGGCTA
ATAAAGCCGTAA

SEQ ID NO: 132, Proteína - *Arabidopsis thaliana*

MKEKVEKLETDDLKWTERLPECPVYRPTKEEFEDPLTYLQKIFPEASKYGICKIVSPLTATVPAGA
VLMKEKSNFKFTTRVQPLRLAEWDSDDKVTFFMSGRTYTFRDYEKMANKVFARRYCSGGSLPDSFL
EKEFWKEIACGKTETVEYACDVDGSAFSSAPGDPLGSSKWNLNKVSRLPKSTLRLLETSIPGVTEP
MLYIGMLFSMFAWHVEDHYLYSINYQHCGASKTWYGIPGSAALKFEKVVKECVYNDDILSTNGEDG
AFDVLLGKTTIFPPKTLLDHNVPVYKAVQKPGEFVVTFPRAYHAGFSHGFNCGEAVNFAMGDWFPF
GAIASCRYAHLNRVPLLPHEELICKEAMLLNSSSKSENLDLTPTELSGQRSIKTAFVHLIRFLHLA
RWSLMKSGLCTGLVSNTYGTIVCSLCKRDCYLAFINCECYSHPVCLRHDVKKLDLPCGTTHTLYLR

FIGURA 10 CONT.

DNIEDMEAAAMKFEKEDGVSDLITTDEDLYKYPSSITLPAAKEDGYTPYSTIYFDFYTEVEMTSHD
QLQSGNPVMSYEANASCISSVADDYECSDYQVNRRANCSSSSDSKLSEEVACSSSKKTRFFPVVQD
EQLVADQESDGSDSECFRVKRRSSLKFENRTVVLDTRESDHHQELKRLKKSHHHEGRYSSSSSVSR
QEEEEDELVISNRKETQQQSDVKMQKKRIENHFGGFKRLKVKGLIKP

SEQ ID NO: 133, DNA - *Arabidopsis thaliana*

ATGGGAATACAGATTATAGGTCAAATTGAGAGAATCAATGGTAAAGAGCTGAGTTACGGCGATTTT
GCAGAGAGATACTTAGCTAAGAACCAGCCAGTGGTAATATCCGATCTCACTGAAGATTGGCGAGCT
CGGGAAGATTGGGTCAGTGAGAATGGAAACCCTAATCTCCACGTATTTGCCACTCACTTTGGAAAA
TCCAGAGTTCAGGTTGCAGATTGTGATACGAGAGAATTTACAGATCAGAAAAGATTAGAAATGTCT
GTTACTGAATTTGTGGAGCAATGGACTAATAAGGATTCTATTGAGGAATCTGTTTTGTATCTGAAA
GATTGGCACTTTGTGAAGGAGTATCCAGACTATACAGCTTACCAAACACCGCCGTTGTTTTCTGAT
GATTGGCTTAACGTGTATCTAGACAATTACCAAATGCATGAGGATAGAGATAGTTTCCAGAAGTAT
GATCAAATAAGCTGCTCTGATTATCGGTTTGTTTATATGGGTGGGAAAGGATCTTGGACACCTCTA
CATGCTGATGTATTTAGATCTTATAGTTGGTCAGCTAATGTCTGTGGTAAAAAAAGATGGCTTTTC
CTTCCTCCTCCTCAAAGTCATCTTGTTTATGACAGGTACATGAAAAATTGCGTTTATGACATTTTT
GAAGAGGTGAATGAAACTAAATTCCCTGGTTTTAAGAAGACCACCTGGCTTGAGTGCATTCAAGAA
CCCGGTGAAATTATCTTTGTTCCCAGTGGATGGCATCATCAAGTATATAACTTGGAAGACACAATA
TCTATTAACCACAATTGGCTCAACGCATACAACCTATCTTGGGTGTGGGATCTACTGTGGAAGGAC
TACAAGGACACTGAAGAGTCAATAGAAGACATAAGAGACATATGTGATGATTTTGAAGCTATCTGC
CAACGTAATCTTGCTGCCAACACAGGAATGAATTTAAATGATTTCTTCCTCTTCATGTCACGGTTC
TCTTTGGGGAACATGGTTCTACTGCAATCTTACTCTGACAAACACAAAAACCTGAATTCGTGTTCA
TTAGCAATGGCCCAAAACCTGTTAATGAATCTCTCAACTATACTGAAAGTCATGATGAAGATGATA
TCTGCAGGCGGTGTAACTGCAGAGGAAGTGTATCTGGACTTGCGAGAAACACTGGAGGATCCACAG
TTTCTCAGATTTGTCAGAGACATGGGAAGAACTTATGCAAGGATTCACATGGAGGAAGAGGATCAG
TTTCTTTCTTCGAAAGAATTATTACAAAAGCTCAGTGGTCTTGCAGGTCCAAATATGCAAATATGT
TCTCCAAAGGATCTTGTTGAAATGATAAACCATCACAATACTTTTTCCTCCCAAATCTATTTTATT
TAG

SEQ ID NO: 134, Proteína - *Arabidopsis thaliana*

MGIQIIGQIERINGKELSYGDFAERYLAKNQPVVISDLTEDWRAREDWVSENGNPNLHVFATHFGK
SRVQVADCDTREFTDQKRLEMSVTEFVEQWTNKDSIEESVLYLKDWHFVKEYPDYTAYQTPPLFSD
DWLNVYLDNYQMHEDRDSFQKYDQISCSDYRFVYMGGKGSWTPLHADVFRSYSWSANVCGKKRWLF
LPPPQSHLVYDRYMKNCVYDIFEEVNETKFPGFKKTTWLECIQEPGEIIFVPSGWHHQVYNLEDTI
SINHNWLNAYNLSWVWDLLWKDYKDTEESIEDIRDICDDFEAICQRNLAANTGMNLNDFFLFMSRF
SLGNMVLLQSYSDKHKNLNSCSLAMAQNLLMNLSTILKVMMKMISAGGVTAEEVYLDLRETLEDPQ
FLRFVRDMGRTYARIHMEEEDQFLSSKELLQKLSGLAGPNMQICSPKDLVEMINHHNTFSSQIYFI

SEQ ID NO: 135, DNA - *Oryza sativa*

ATGTCGATGCGGGAGTTTGTCGACCACTGGGCTGCGAGTTCGAGCAATGGGGATTCTGATGGTTCC
CTGCTGTACCTGAAAGATTGGCATTTTGTGAAGGAGTACCCTGGTTATGTTGCATACACCACACCA
ACATTCTTTGCCGATGATTGGCTCAACATGTATCTTGATAGTCATCCTATACATCGAGATTCTGAC
ATTGCCAACCATACAAATGAAATAAACTGTGCTGACTATCGGTTTGTTTACATGGGACCAAAAGGA
ACTTGGACTCCTCTTCATGCTGATGTTTTTAGGTCATACAGCTGGTCGGCAAATGTATGCGGCAGA
AAACTATGGCTCTTTCTACCACCATCACAAAGTCATTTTGTATTTGATAGGAACCTACGATCTTCG
GTCTATAACATAAATGATGATGTTTCTGAAAAGCAGTTTCCTGAATTCAATAATACTAAATGGCTA
GAGTGCACTCAGGAGCAGAATGAAATCATTTTTGTGCCTAGTGGGTGGTATCACCAAGTCCATAAC
CTGGAGGATACTATATCAATAAACCATAATTGGTTTAATGGCTACAACCTTCATTGGGTGTGGAAC

FIGURA 10 CONT.

TTGCTTCATGAAGATTACAAGGTAGCAAAAGATTATATTGAAGACATTCGGGACATATGTGACGAT
TTTGAGGGACTATGTCAACGCAACTTAGCAGCAAATACAGGCATGAACTTCTATGACTTCTTCGTC
TTCATAACACGTTTTGCTTTAGCCAATATAGTTGAGCTTTACCATATTCAGAATCCAAAAGATACT
GACTTCATTTCAGCCGAAACTGCTAACCACTTTGTCTATAATCTCATGTCGATCCGTGATGTTGCA
TCAAAAATGGTATCTACAGAAGCTTTCAATACCGAGAATATTTGTAATATCTCAGAACAAAACCGG
AGTGCTTTCTCTGATATCATAAAAATATTGGAAGAAGAGAGTTTCAGAAGGTTATTGGTGGCGTTG
TCAAAGGCATATAATTACATAGACAGAGGACAAAAAGATTGTCTCAAAATGAAGGATTCAAGTCAG
AAGGGTTGTTTGTCAGTAACCTGCTTAAAACCTGACTGTAATGTTGTTGGTGACATTATTTCTTTC
ATGCGTGAGATTCATGGACCTATGGATTTGGTAACACTAATTGATAGTGCTCTCTCTGATAGATAG

SEQ ID NO: 136, Proteína - *Oryza sativa*

MSMREFVDHWAASSSNGDSDGSLLYLKDWHFVKEYPGYVAYTTPTFFADDWLNMYLDSHPIHRDSD
IANHTNEINCADYRFVYMGPKGTWTPLHADVFRSYSWSANVCGRKLWLFLPPSQSHFVFDRNLRSS
VYNINDDVSEKQFPEFNNTKWLECTQEQNEIIFVPSGWYHQVHNLEDTISINHNWFNGYNLHWVWN
LLHEDYKVAKDYIEDIRDICDDFEGLCQRNLAANTGMNFYDFFVFITRFALANIVELYHIQNPKDT
DFISAETANHFVYNLMSIRDVASKMVSTEAFNTENICNISEQNRSAFSDIIKILEEESFRRLLVAL
SKAYNYIDRGQKDCLKMKDSSQKGCLSVTCLKPDCNVVGDIISFMREIHGPMDLVTLIDSALSDR

SEQ ID NO: 137, DNA - *Oryza sativa*

ATGAGGCCCTCGCCGCCTCCCGCCGCCCCGGCGGCGGAACCGGTGCCCCCGTGGCTGAGATCGCTG
CCCGTCGCGCCCGAGTTCCGCCCCACCGCCGCCGAGTTCGCCGATCCCGTCTCCTACATCCTCAAG
ATCGAGCCCGCCGCGGCGCCGTACGGCATCTGCAAGGTCGTGCCGCCGTTGCCCCCGCCGCCGAAG
AAGGCCACCTTCTCCAACCTCTCCCGCTCCTTCGCCGCGCTCCACCCGGACGACCGCTCGCCGTCC
TTCCCCACCCGCCACCAGCAGGTCGGCCTCTGCCCGCGCCGCACGCGCCCGGGGCTCAAGCCCGTG
TGGCGGTCCTCCCACCGCTACACGCTCCCGCAGTTCGAGTCCAAGGCCGGCGCCACCCGCAAGTCG
CTCCTCGCCGGCCTCAACTTCCCTGCCTCCAGGCAACTCACCCCGCTCGACCACGAGGTGCTCTTC
TGGCGCGCATCGGCCGACCGCCCCATCGTGGTCGAGTATGGCAGCGACATGTCGGGCTCGGGATTC
TCCCCTTGCGCCGCGCAGCCGCAGCCGCCGCCGCAGCAGCAACCGACGGCGCGGGCGGCGGCGCAC
CTTGGGGAGACGGCGTGGAACATGCGCGGCGTGGCTAGGAGCCCCGGATCGCTGCTGCGGTTCATG
CCGGAGGATGTCCCCGGGGTCACCACGCCGATGCTCTACGTCGGGATGATGTTCAGTTGGTTCGCG
TGGCACGTCGAGGACCACGACCTGCACAGCCTCAACTACATGCACCTCGGGGCCGCCAAGACGTGG
TACGGCGTGCCGCGCGACGCCGCGCTCGCGTTCGAGGACGTGGTGCGCGAACATGGCTACGGCGGG
GAGGTGAACCCCCTAGAAACCTTCGCTACATTGGGGCAGAAAACAACAGTGATGTCCCCTGAAGTG
CTTGTGGAGTCGGGTATCCCCTGCTGCAGATTGGTGCAGAATGCAGGGGAATTTGTGGTCACTTTT
CCAGGATCTTATCACTGCGGTTTCAGTCATGGATTCAACTGTGGGGAAGCATCAAATATAGCTACT
CCTGAATGGTTGAGAATAGCAAAAGAGGCAGCAATCCGAAGAGCTTCAATCAATCGTCCTCCCATG
GTTTCGCACTATCAATTACTTTACGATCTCGCACTGTCTATGCGTTTTAGGGAACCATCCAATGGT
GAAATGGAGACACGGAGCTCTCGAATAAAGGAGAAGAAGAAATGCGAAGGGGAACAACTGGTCAAG
AAGATGTTCATCCAAAATGTTATTGAAGATAATGAACTGCTTAGTCATCTTCTAAATGACGGATCA
TCTTGTATTATTCTTCCAGCAAATGCCCATGATGGTCCTGGTCTTTCTACTTTGCGCTCAACAGAT
CAATCAAATATGAATTCCAGAATATCACATAACCTATGCAGTAGGGAAGAAGCTCCAGAAGCCTCA
GGATGCTTATCGCCGAACAGGAATGGTGATACCAGAAATTGTATTTCATCAGATACGCATAACATG
GAAGGTGATAAGGGAGATATAATGAGCGCTACTGGATTGTTAGATCAGGGTCTATTATCATGTGTC
ACTTGTGGAATTTTAAGTTTTTCATGTGTGGCTGTACTCAAACCAAGAGATTCTACAGCACGATAC
TTGATGTCTGCTGATTCTAATTCAATTAATAACCAGCTTTCTATTTCTGGAGGGAGTATTCTGGCA
GATGCGCCAACAAATGAAAGGAATGGTGTCATTTCTCGTCCATATTCTGAACACTGTTGCAACGAA
ATAATGGCCGATGATGCAGAAATCGATAAGAATTCTGCTCTTGATCTCTTGGCATTTGCTCATGGG
GGTCAACCAGATCCTGAAGAAGATCCACTGGAGAAAATACTGAAGATTGCACATGGCATTAACAAA

FIGURA 10 CONT.

TCACAACCTAATAGTTCAAATAATGTTGGTTGTGTTGGAACAAAGCTGTCCTCAAGTAGCACAGAG
CGCCAAGAAAGACCATCTTCCCAGAACGCTCATTGTAATGGTAGCTCAGTTATTTCAAATGGACCG
AAAGGTGTTCGCACGAGAAATAAGTATCAACTTAAGATGGTACTTTCTGAAGGTTTTCAGGCAAAG
GATATCTATTCTGCGAAGGAAAAGAAGGTTCAATCAGAACCATCAAGCTCAAAAGGGGATGTTAAG
GAGACAATTGATGTTAGTGGCACAGAGAACGATGTCGGATGTAAAAGCACTACAATCTCTGTCAGT
GAACACAGGGGTTCTACAAAGAATATGTATTCAGTGAAGGAAAAGAAGGTTCAATCGAAACCATCA
AGCTTAAAGGGGACTGTTAAAGAGACAGTTGATGTCAGTGGCACAGAGAACGATGCTAGATGTAAA
AGCATTACAATCTCTGTCAGTGAACACAGGGGTTCTACACCTATGACCAATAGTTTAGCTGCATCA
ATCGTGAAACCTGACAAGGATTCATCAAGAATGCATGTATTTTGTCTTGAGCATGCTATTGAGGTT
GAGAAGCAACTGCATGCTATAGGTGGTTCTAATATTATGCTTATATGTCGTCCAGAATATCCCAAA
ATAGAAGCGGAAGCAAGATTGTTGGGTGAAGAAATGGGACTGGTATATGACTGGAAGGGCATTCAT
TTCAAGGAGGCCAACATGGAGGACAGGCAGAAGATACAGGAAGTTCTGCGGGATGAGGAAGCAATA
CCTACAAGCAGTGATTGGGCTGTAAAATTGGGTATTAACCTTTATTATAGTGCAAATCTGGCCAAA
TCTCCTTTATATAACAAACAAATGCCATACAACAGAGTCATTTATAGGGCATTTGGCTGTGATTCT
CCTAATGACTCGCCAGTAATGTTCAATACTTGCGAAAGGAAACAAAGCCACCAGAAGAAAATAGTG
GTTGCTGGCCGGTGGTGTGGGAAAGTATGGATGTCAAAACAGGTCCATCCATACTTGGCTCACAGA
GTTGAAAGCCAGGAAGCGGAAGAAGCAGATAGAATCTGTTCTTATCATTTTGATGAGAAGCACAAA
GCTGAGCCAGTTGGGAATTCTAGTAGAGTGGAAGCCTCCAAAAGAAAGAGTAGCAGTTTGACAGAC
GTAACAGAATCAAGCAATAGAAGGGGGGAAATACCTGGTGAGGAAACAAACACCAAGAGGCCCAAA
CATTCTCAAGAGAACAATCTCAGAGCTTTGGAAACCGCTGCTGAAGTTGTGGTGCCCTCACCAGCT
GGAACAGGTCTCCGCGTTAGCTCCAGGATTGCCAACAGGGCAAACAAGCTAAAATCAAAGATGGAA
AAGGAGGATGTCCCTTCTAGCCGTCCAAAGTCTAACATAAAGGAGAAAAGCAGTCATGCTAGCGGT
CAGAAATCAAATGTTCAAGAGGCGAATGCTAATTCTGCTAGCCATTTGAGAGCAATGCCACCAAAA
CAGAAGGCTGAAGCAGAAGCAAAGAAGCAAATAAGAACTCCAAAACCTCCAAAGCAAGCAGTGGAG
TACTCGTGCGATATTGAGGGTTGTTCTATGAGTTTTCGTACAAAGCGGGACTTGTCTCTGCACAAG
AGTGATATCTGCCCAGTGAAAGGCTGCGGGAAGAAGTTCTTCTCGCACAAGTATCTGCTGCAGCAT
CGCAAGGTTCACACAGATGACCGTCCGCTAACATGCCCATGGAAGGGCTGCAACATGGCATTCAAG
TGGCCATGGGCAAGGACTGAACATCTCAGGGTTCATACAGGCGACAGGCCCTATGTTTGCCATGAG
CCAGGTTGCGCACAGACATTCAGGTTTGTCTCAGATTTCAGCCGTCACAAGCGCAAGACAGGTCAT
TCTGTCAAGAAGAAGAAGAAAGCAAAGAGTTGA

SEQ ID NO: 138, Proteína - ***Oryza sativa***

MRPSPPPAAPAAEPVPPWLRSLPVAPEFRPTAAEFADPVSYILKIEPAAAPYGICKVVPPLPPPPK
KATFSNLSRSFAALHPDDRSPSFPTRHQQVGLCPRRTRPGLKPVWRSSHRYTLPQFESKAGATRKS
LLAGLNFPASRQLTPLDHEVLFWRASADRPIVVEYGSDMSGSGFSPCAAQPQPPPQQQPTARAAAH
LGETAWNMRGVARSPGSLLRFMPEDVPGVTTPMLYVGMMFSWFAWHVEDHDLHSLNYMHLGAAKTW
YGVPRDAALAFEDVVREHGYGGEVNPLETFATLGQKTTVMSPEVLVESGIPCCRLVQNAGEFVVTF
PGSYHCGFSHGFNCGEASNIATPEWLRIAKEAAIRRASINRPPMVSHYQLLYDLALSMRFREPSNG
EMETRSSRIKEKKKCEGEQLVKKMFIQNVIEDNELLSHLLNDGSSCIILPANAHDGPGLSTLRSTD
QSNMNSRISHNLCSREEAPEASGCLSPNRNGDTRNCISSDTHNMEGDKGDIMSATGLLDQGLLSCV
TCGILSFSCVAVLKPRDSTARYLMSADSNSINNQLSISGGSILADAPTNERNGVISRPYSEHCCNE
IMADDAEIDKNSALDLLAFAHGGQPDPEEDPLEKILKIAHGINKSQPNSSNNVGCVGTKLSSSSTE
RQERPSSQNAHCNGSSVISNGPKGVRTRNKYQLKMVLSEGFQAKDIYSAKEKKVQSEPSSSKGDVK
ETIDVSGTENDVGCKSTTISVSEHRGSTKNMYSVKEKKVQSKPSSLKGTVKETVDVSGTENDARCK
SITISVSEHRGSTPMTNSLAASIVKPDKDSSRMHVFCLEHAIEVEKQLHAIGGSNIMLICRPEYPK
IEAEARLLGEEMGLVYDWKGIHFKEANMEDRQKIQEVLRDEEAIPTSSDWAVKLGINLYYSANLAK
SPLYNKQMPYNRVIYRAFGCDSPNDSPVMFNTCERKQSHQKKIVVAGRWCGKVWMSKQVHPYLAHR
VESQEAEEADRICSYHFDEKHKAEPVGNSSRVEASKRKSSSLTDVTESSNRRGEIPGEETNTKRPK

FIGURA 10 CONT.

HSQENNLRALETAAEVVVPSPAGTGLRVSSRIANRANKLKSKMEKEDVPSSRPKSNIKEKSSHASG
QKSNVQEANANSASHLRAMPPKQKAEAEAKKQIRTPKPPKQAVEYSCDIEGCSMSFRTKRDLSLHK
SDICPVKGCGKKFFSHKYLLQHRKVHTDDRPLTCPWKGCNMAFKWPWARTEHLRVHTGDRPYVCHE
PGCAQTFRFVSDFSRHKRKTGHSVKKKKKAKS

SEQ ID NO: 139, DNA - *Oryza sativa*

ATGCCGCCCAAGAGGAAGCGCGGTGGCGGGAGGCCGCGCAAGGCCCCGGAGGACGCCGCCGCCAAG
GAAAACGGTGAGAAGACAAACAAGGAGGAGGAAACGCAGGCTTCTCCTGAGGAGAATGGCGCAGGG
CAGACTCAGGCCTCGAGAACCGCAAGAAAGCGCAGGAAGGGTCCTGTTGCTGATCCCTCTTCGACT
GAACTTCCACCCAGGAAGCTCAGGGACAGGCGGAACGTTCCAGCGGTAGACTACAAGGAGAATAAG
CACACAAAAAAGATGGATGGAACTTCAACCATGTGCCATCAGTGCCAAAGGAAGGACAGTGGAAGA
GTTGTGAGGTGCCGGAATGGTGCTGAAAAAAACAGGAGGCACAGATATTGTGTCAAGTGCATCAAA
CGCTGGTACCCGCATCTAACAGAGGACGATTTTGAGAACTGCTGTCCAGTTTGTCACAATAACTGC
AATTGTAAGACTTGTCTGCGGACAAATGTAATAAATAAGGGTGACAAGGAGTTTGCCGATGGTAAG
AACAAAATTAAATACTCTCTGCGTATTGCACGCTTTTTGCTCCCTTGGTTGAAACAACTTCACCAA
GAGCAGATGCTAGAGAAAAGTGTTGAGGCCACAATTAAAGGGATCGATGTAACTGACTTGGAGGTT
CCTCAAGCTCAGTTTAATAATGATGAGCGGATTTACTGTGATAACTGCAGAACATCCATTGTTGAC
TTCCACAGAAGTTGTAAAAGTGGTCACTATGATCTCTGCCTCAGCTGCTGCCAGGAGTTGCGTCAA
GGTCTTACTACTGGTACTGTTGTCACTTGTGACACAGCTGTTGATGTACCTGAAATAGAAGGCAAG
GAAGGTTTGCAAGAGGGAAGTAGTCATAGCAGTGCTGTAGGTCAAGGGGCTTCTGATCAACAGAAT
GACAGGTTGATAGGTAGTGCAGCTCCTTCAGAGGATTGCACTCCTAGTTTGATATGGAGGGCAAAA
AGCAACGGAAGCATACCATGCCCACCTAATGCAGGTGGTTGCGGGGATTGTCTTCTCGAACTTAGG
TGTTTGTTTAAGGAGAACTTTATTTCTGATTTACTGGATAAAGTCAATTCAGTGGTCAATAAGGAA
ACAGAGCAGGAGTTGGGAGGCTCTAGATGTTCCTGCTTCACTGAATCCGGTGAAGTGAACAATGAA
ACATCACGGAAATCAGCTTGCAGAGAGGACTCCAATGATAACTACATATATTGTCCAACTGCTAGA
GAAGTTCAAAGTGGAGCTTTAGATCATTTTCAGCAGCATTGGTTGAATGGTCAGCCAGTCATTGTT
CGTGACGTGCTTGAGCTAACTTCTGGACTGAGCTGGGAGCCAATGGTTATGTGGCGGGCCTTACGG
GAAAAGAGAGACAAGAAAGAGCATGAACGGCTCTCAGTTATAGCTCTTGATTGTCTGACATGGTTT
GAGTTTATGTACCACCAGGATGTGTTGGTGGTGCCAGTTTCATATCTGGGGTTCAACAGCACCATT
GAAACTACTTTATATTTCAAACTGGTTGATATAAATATTCACATGTTTTTTGAGGGGTATTCCCGT
GGAGCTGTTGGTTCGGAGGATTTGCCTGTGCTACTTAAGCTTAAAGATTGGCCACAACATAGTTCC
TTTGAGGAGCGCCTGCCGCGGCATGGTGCTGAGTTCATGTCTGCACTACCATTTCGTGAGTATACA
GATCCAAAATCTGGTCCCCTTAATCTGGCAGTGAAGCTTCCAAAACATGTTAAAAAGCCAGATCTT
GGTCCAAAGACATACATTGCTTATGGTGTCGCCCAAGAGTTGGGAATAGGTGATTCTGTTACGAAG
ATCCATTGCGACATGTCTGATGCGGTTAATATTCTCATGCATACTGATGAGGTAGAACTCAAAGCT
GAAAGGATTACAGCAATAGAGAAAAAGAAAGAGAGCTTACGTAAAGATGGCAAGAATCTTCATGTT
CTGCGACCAGACCATGACGATGATACATCAATAGCTCTCAGTGAATCAACTGAAGTGCCAAGATCT
CGAGGACTTGAAAATGGCTCAAGCATTAAGCAGCCAGCACCAAATGTTGCTGTAATGGACCAGGGA
GGTGTTCATACAGATATGGTAGCTGATGAAGCTGAGGGAAACTTGAGTCTGTCTAATGGGCAATCA
CCCAATCAAAGTGATGCACATAATATGGATATTACTTTCTCCAAGGGGGAAACAGACCATTCTATC
TGCACAATTAATGGGGGAGAGGAGATGGGCAATGGCTTTGGGAGGGAAGATAAATGTAAATCTTCT
CATGGTGTAGGATCATCTGAATCTTCTGATTGTCAAAGACGCAGTAGGCGACGTGATGCTTGCTCT
TCCAGTGCAACAGGGGAAATAAACGAAACCAGCATGGAGACTAACAAATTTACCATATCTATTGAG
CCCAAAGATGATCACCCATTTGTTGAGGGAAATCAGACAGAGGGTGGTGCATTGTGGGATATCTTC
CGACGTGAAGATGTCAGCAAATTGCATGATTATCTGATGAAGCATGCAGAAGAGTTTAGACATTAT
AATTATGAAACAGTCAAACAGGTTTCTCATCCTATTCACGATCAATGCTTTTATCTAACAAATGAG
CACAAGAGAAAGCTCAAGGAAGAACATGGAATTGAGCCTTGGACATTTGAACAGAAACTTGGTGAG
GCTGTGTTTATCCCAGCAGGATGTCCCCACCAAGTCAGAAATTTGAAGTCATGTATAAAGGTTGCA

FIGURA 10 CONT.

CTTGACTTTGTTTCTCCGGAAAATGTGCAAGAGTGCATCAGGCTGACAGAAGAATTTCGCTTGCTT
CCGAAGGGTCATAGGGTGAATGAAGATAAACTAGAGGTTAAGAAGATAGCTCTATATGCACTTGAT
CAAGCCATTGACGATATCACAGGCAAAAGTTGCAATGAAAGGACAAAGGATGAAGGGGAAGAGGAG
GCGTCTGCCCCAAGTGTTAGTTAA

SEQ ID NO: 140, Proteína - ***Oryza sativa***

MPPKRKRGGGRPRKAPEDAAAKENGEKTNKEEETQASPEENGAGQTQASRTARKRRKGPVADPSST
ELPPRKLRDRRNVPAVDYKENKHTKKMDGTSTMCHQCQRKDSGRVVRCRNGAEKNRRHRYCVKCIK
RWYPHLTEDDFENCCPVCHNNCNCKTCLRTNVINKGDKEFADGKNKIKYSLRIARFLLPWLKQLHQ
EQMLEKSVEATIKGIDVTDLEVPQAQFNNDERIYCDNCRTSIVDFHRSCKSGHYDLCLSCCQELRQ
GLTTGTVVTCDTAVDVPEIEGKEGLQEGSSHSSAVGQGASDQQNDRLIGSAAPSEDCTPSLIWRAK
SNGSIPCPPNAGGCGDCLLELRCLFKENFISDLLDKVNSVVNKETEQELGGSRCSCFTESGEVNNE
TSRKSACREDSNDNYIYCPTAREVQSGALDHFQQHWLNGQPVIVRDVLELTSGLSWEPMVMWRALR
EKRDKKEHERLSVIALDCLTWFEFMYHQDVLVVPVSYLGFNSTIETTLYFKLVDINIHMFFEGYSR
GAVGSEDLPVLLKLKDWPQHSSFEERLPRHGAEFMSALPFREYTDPKSGPLNLAVKLPKHVKKPDL
GPKTYIAYGVAQELGIGDSVTKIHCDMSDAVNILMHTDEVELKAERITAIEKKKESLRKDGKNLHV
LRPDHDDDTSIALSESTEVPRSRGLENGSSIKQPAPNVAVMDQGGVHTDMVADEAEGNLSLSNGQS
PNQSDAHNMDITFSKGETDHSICTINGGEEMGNGFGREDKCKSSHGVGSSESSDCQRRSRRRDACS
SSATGEINETSMETNKFTISIEPKDDHPFVEGNQTEGGALWDIFRREDVSKLHDYLMKHAEEFRHY
NYETVKQVSHPIHDQCFYLTNEHKRKLKEEHGIEPWTFEQKLGEAVFIPAGCPHQVRNLKSCIKVA
LDFVSPENVQECIRLTEEFRLLPKGHRVNEDKLEVKKIALYALDQAIDDITGKSCNERTKDEGEEE
ASAPSVS

SEQ ID NO: 141, DNA - *Oryza sativa*

ATGGCGGCGGAGGAGGTGGAGGCGGCGGCGATCGACGGATCCGGCGAGGAGGCGAAGCGGAAGAGT
GGTAAGCAGAGGGGATCGGGCGCCAAGGGGAGGCGCCGGAACGGCGACCGGGCTTTCCGCCCGCCG
GCGATGAGGCCGGAGGAAGAAGGGCGCGGTGTTGCGACGAGGCCCGGGGCGCTCAGGGAAAGGAAG
CCGCCGCCCAACGCCTTCAACGCGCCAGACGATGATGAGGATGTCGAGAAAACTGATCAACTTCTC
GAGCCTCTCAACAAGCCCAAAAGACGTGATGCAGGGAAGAAGAGAGGACCCAGAAAAAAGAAAGTG
GATCAAGAGAACATCAAAACCCACAGGCATAACGCAAATGCAGTAAAAGGCAAGATGTTAGTCAAT
GATAAGGTTTCGAAGACTGAAAAGAAGCGTAAGAGAGGAGACACAGGAGCGGCAGAAAATAATGGG
AAAGGGAAGAAAATGTTGACAGGTGAAAACGCCTTAATGTGTCACCAATGCCAGAGAAATGACAAA
GGGAGAGTGGTCTGGTGCAAGACTTGCAACAATAAGCGCTTTTGTGTGCCATGCATTAATCAATGG
TATCCTGATTTGCCTGAAAATGAATTTGCTGCAAAATGCCCTTATTGTCGCAAGAATTGTAATTGC
AAGGCATGCCTGCGAATGAGAGGTGTCGAAGAGCCTCCAAGAAAGGAAATTTCTAAAGAAAATCAG
ATTCGTTATGCCTGTCATGTGTTACGCTTGTTGCGTCCTTGGCTGATAGAACTGCGACAGGAGCAG
ATGGCAGAGAAGGAATTAGAGGCTAAGATACAAGGTGTTTCAGTTGATCAAATAAAGGTGGAACAA
GCTGTGTGTGATCTAGATGAGCGTGTCTATTGCAATAGATGCAGCACTTCTATAGTTGATTTTCAT
AGAAGCTGCAAGCACTGTTTCTATGACTTGTGCTTGACCTGCTGCCAGGAGCTTCGCAAAGGTGAG
ATCCCTGGAGGAGAAGAGGTGGAAATTTTGGATCCTGAAGAAAGAGACAAGGATTATGCTTTTGGT
AAGATTCTGTCAGATGGTGAGAACCAAAGGGACTCTTTGAAGTGTCGCAGCGACACCCAAAACAGT
GAGTCAAACAAAGGCATGGCTTCTGACGAAAACCAAAAGAAGGCTCTGTTACTTTGGAAAGCAAAC
AGTAATGGTAGCATACCTTGCCCACGAAAGGAAAAGGAAGACTGCAGTTTTTCATCTCTAGATCTG
AAATGCTTGTTTCCAGAGAAGTTGCTCCCTGAATTAGAAGATAGATCTGAGAAGGTCTTCTGGAGT
GAAACATTTGCAAAAGAACTAGGTAGAACAAGTGAGCTGTGCCCTTGTTTTGATCACTCAGGCAAG
ATAAGAAGTGACAGCAAGAAATTACGTCAGGCTGCAAATAGGGAGGACTCGAGTGATAACTACTTA
TACTGTCCGGTGGCCACTGATATCCAGGACGCTGACCTGTTGCACTTTCAGATGCATTGGGCAAAG
GGCGAACCAGTTGTTGTTTCAGACACTCTTAAGCTAACTTCTGGTCTTAGCTGGGAGCCGATGGTT

FIGURA 10 CONT.

ATGTGGCGGGCAGTACGAGAACGAACCAAAGGGAAGGCAGAAGATGAGCAGTTTGCTGTGAGGGCA
GTAGATTGCCTCGACTGGTGTGAGGTGGAAATTAACATACACATGTTTTTTATGGGATATACAAGG
GGCAGAACTCATCCCAGGACCTACTGGCCTGAGATGCTTAAGCTAAAGGACTGGCCCCCTTCTAGC
TCATTTGATCAGCGATTGCCTCGCCATGGTGCTGAATTTATCAGTGCATTGCCATTTCCGGAGTAT
ACTGATCCACGATATGGTCCGCTAAATCTTGCAGTGAAGCTTCCTGGTGGTGTCCTGAAGCCAGAT
CTTGGGCCAAAAACTTACATTGCCTATGGATTTTCTGAAGAGTTGGGCAGGGGTGACTCTGTAACC
AAACTTCACTGTGACATGTCTGATGCAGTAAACATCTTAACACACACAGCGGAAGTGCCCTGTGAG
ACTTATGATGCTGTTCAGATTAAAAACACGCAGAAAAAGATGAAAATGCAAGATGACATGGAAATA
TATGGGATGATAGAATCAGGCAGTGAACTCAAGCCATCAGCATGCCCAGTTGAATTGGGGAACAAA
GCTGTTGGTGAAGCACCAAAGGCCTCATGCAGCAAAGAAAATGTTCATACCTTGAAAGACAAGTCT
AATGGTTTGGATATAAACGCTTCACCACCTGATGATGCTGGAGGTGATGCCAGGGATGAAGCATTG
TCCTATGAATCTGTGGTACATAGTGATGTAGCTCAGTGTCCAAACCATAACCATGAGACCAATAAT
TCTGATGATGCACGTATTGGAGCTCAGAGATGTCAAAAGAAAGCTAAAGGTCGTCCTCCGAAGACA
GGTTCAGGGGTATCAGAGCACCAGGAAAGTGGTGGTGCTTTGTGGGATATTTTCCGAAGAGAGGAC
TCTGAGAAATTACAAGACTTCCTTAGGAAGCATGCGCCAGAATTTCGGCACATCCACTGCAATCCA
GTGAAACAGGTTATTCATCCAATTCATGATCAAGCTTTCTATTTAACTGCGGAGCATAAAAGAAAG
CTGAAGGAAGAATATGGTGTTGAACCTTGGACCTTTGAACAAAAGCTTGGCGAAGCAGTTTTGATT
CCTGCCGGATGTCCTCACCAAGTCAGGAATTTGAAGTCCTGCATCAAGGTTGCACTGGACTTTGTT
TCCCCCGAGAATGTTGGTGAGTGTGTTAGGCTGACCAAGGAATTTAGACGACTTCCATCTTCCCAC
AGGGCTAAAGAAGATAAGCTAGAGATCAAGAAGATGGCTTTCCATGCTCTGAATGAAGTTCTAAAC
TTTCTGGATCCTCCTTCCTCAGAAGGGTCGAAAGAAGCGGCGGAGAAGCCTCGAAGGGGGCGTGGT
CGCCCAAGAAAACACTAG

SEQ ID NO: 142 Proteína - *Oryza sativa*

MAAEEVEAAAIDGSGEEAKRKSGKQRGSGAKGRRRNGDRAFRPPAMRPEEEGRGVATRPGALRERK
PPPNAFNAPDDDEDVEKTDQLLEPLNKPKRRDAGKKRGPRKKKVDQENIKTHRHNANAVKGKMLVN
DKVSKTEKKRKRGDTGAAENNGKGKKMLTGENALMCHQCQRNDKGRVVWCKTCNNKRFCVPCINQW
YPDLPENEFAAKCPYCRKNCNCKACLRMRGVEEPPRKEISKENQIRYACHVLRLLRPWLIELRQEQ
MAEKELEAKIQGVSVDQIKVEQAVCDLDERVYCNRCSTSIVDFHRSCKHCFYDLCLTCCQELRKGE
IPGGEEVEILDPEERDKDYAFGKILSDGENQRDSLKCRSDTQNSESNKGMASDENQKKALLLWKAN
SNGSIPCPRKEKEDCSFSSLDLKCLFPEKLLPELEDRSEKVFWSETFAKELGRTSELCPCFDHSGK
IRSDSKKLRQAANREDSSDNYLYCPVATDIQDADLLHFQMHWAKGEPVVVSDTLKLTSGLSWEPMV
MWRAVRERTKGKAEDEQFAVRAVDCLDWCEVEINIHMFFMGYTRGRTHPRTYWPEMLKLKDWPPSS
SFDQRLPRHGAEFISALPFPEYTDPRYGPLNLAVKLPGGVLKPDLGPKTYIAYGFSEELGRGDSVT
KLHCDMSDAVNILTHTAEVPCETYDAVQIKNTQKKMKMQDDMEIYGMIESGSELKPSACPVELGNK
AVGEAPKASCSKENVHTLKDKSNGLDINASPPDDAGGDARDEALSYESVVHSDVAQCPNHNHETNN
SDDARIGAQRCQKKAKGRPPKTGSGVSEHQESGGALWDIFRREDSEKLQDFLRKHAPEFRHIHCNP
VKQVIHPIHDQAFYLTAEHKRKLKEEYGVEPWTFEQKLGEAVLIPAGCPHQVRNLKSCIKVALDFV
SPENVGECVRLTKEFRRLPSSHRAKEDKLEIKKMAFHALNEVLNFLDPPSSEGSKEAAEKPRRGRG
RPRKH

SEQ ID NO: 143, DNA - *Oryza sativa*

ATGTTAGTCAATGATAAGGTTTCGAAGACTGAAAAGAAGCGTAAGAGAGGAGACACAGGAGCGGCA
GAAAATAATGGGAAAGGGAAGAAAATGTTGACAGGTGAAAACGCCTTAATGTGTCACCAATGCCAG
AGAAATGACAAAGGGAGAGTGGTCTGGTGCAAGACTTGCAACAATAAGCGCTTTTGTGTGCCATGC
ATTAATCAATGGTATCCTGATTTGCCTGAAAATGAATTTGCTGCAAAATGCCCTTATTGTCGCAAG
AATTGTAATTGCAAGGCATGCCTGCGAATGAGAGGTGTCGAAGAGCAGCCTCCAAGAAAGGAAATT
TCTAAAGAAAATCAGATTCGTTATGCCTGTCATGTGTTACGCTTGTTGCGTCCTTGGCTGATAGAA

FIGURA 10 CONT.

CTGCGACAGGAGCAGATGGCAGAGAAGGAATTAGAGGCTAAGATACAAGGTGTTTCAGTTGATCAA
ATAAAGGTGGAACAAGCTGTGTGTGATCTAGATGAGCGTGTCTATTGCAATAGATGCAGCACTTCT
ATAGTTGATTTTCATAGAAGCTGCAAGCACTGTTTCTATGACTTGTGCTTGACCTGCTGCCAGGAG
CTTCGCAAAGGTGAGATCCCTGGAGGAGAAGAGGTGGAAATTTTGGATCCTGAAGAAAGAGACAAG
GATTATGCTTTTGGTAAGATTCTGTCAGATGGTGAGAACCAAAGGGACTCTTTGAAGTGTCGCAGC
GACACCCAAAACAGTGAGTCAAACAAAGGCATGGCTTCTGACGAAAACCAAAAGAAGGCTCTGTTA
CTTTGGAAAGCAAACAGTAATGGTAGCATACCTTGCCCACGAAAGGAAAAGGAAGACTGCAGTTTT
TCATCTCTAGATCTGAAATGCTTGTTTCCAGAGAAGTTGCTCCCTGAATTAGAAGATAGATCTGAG
AAGGTCTTCTGGAGTGAAACATTTGCAAAAGAACTAGGTAGAACAAGTGAGCTGTGCCCTTGTTTT
GATCACTCAGGCAAGATAAGAAGTGACAGCAAGAAATTACGTCAGGCTGCAAATAGGGAGGACTCG
AGTGATAACTACTTATACTGTCCGGTGGCCACTGATATCCAGGACGCTGACCTGTTGCACTTTCAG
ATGCATTGGGCAAAGGGCGAACCAGTTGTTGTTTCAGACACTCTTAAGCTAACTTCTGGTCTTAGC
TGGGAGCCGATGGTTATGTGGCGGGCAGTACGAGAACGAACCAAAGGGAAGGCAGAAGATGAGCAG
TTTGCTGTGAGGGCAGTAGATTGCCTCGACTGGTGTGAGGTGGAAATTAACATACACATGTTTTTT
ATGGGATATACAAGGGGCAGAACTCATCCCAGGACCTACTGGCCTGAGATGCTTAAGCTAAAGGAC
TGGCCCCCTTCTAGCTCATTTGATCAGCGATTGCCTCGCCATGGTGCTGAATTTATCAGTGCATTG
CCATTTCCGGAGTATACTGATCCACGATATGGTCCGCTAAATCTTGCAGTGAAGCTTCCTGGTGGT
GTCCTGAAGCCAGATCTTGGGCCAAAAACTTACATTGCCTATGGATTTTCTGAAGAGTTGGGCAGG
GGTGACTCTGTAACCAAACTTCACTGTGACATGTCTGATGCAGTAAACATCTTAACACACACAGCG
GAAGTGCCCTGTGAGACTTATGATGCTGTTCAGATTAAAAACACGCAGAAAAAGATGAAAATGCAA
GATGACATGGAAATATATGGGATGATAGAATCAGGCAGTGAACTCAAGCCATCAGCATGCCCAGTT
GAATTGGGGAACAAAGCTGTTGGTGAAGCACCAAAGGCCTCATGCAGCAAAGAAAATGTTCATACC
TTGAAAGACAAGTCTAATGGTTTGGATATAAACGCTTCACCACCTGATGATGCTGGAGGTGATGCC
AGGGATGAAGCATTGTCCTATGAATCTGTGGTACATAGTGATGTAGCTCAGTGTCCAAACCATAAC
CATGAGACCAATAATTCTGATGATGCACGTATTGGAGCTCAGAGATGTCAAAAGAAAGCTAAAGGT
CGTCCTCCGAAGACAGGTTCAGGGGTATCAGAGCACCAGGAAAGTGGTGGTGCTTTGTGGGATATT
TTCCGAAGAGAGGACTCTGAGAAATTACAAGACTTCCTTAGGAAGCATGCGCCAGAATTTCGGCAC
ATCCACTGCAATCCAGTGAAACAGGTTATTCATCCAATTCATGATCAAGCTTTCTATTTAACTGCG
GAGCATAAAAGAAAGCTGAAGGAAGAATATGGTGTTGAACCTTGGACCTTTGAACAAAAGCTTGGC
GAAGCAGTTTTGATTCCTGCCGGATGTCCTCACCAAGTCAGGAATTTGAAGTCCTGCATCAAGGTT
GCACTGGACTTTGTTTCCCCCGAGAATGTTGGTGAGTGTGTTAGGCTGACCAAGGAATTTAGACGA
CTTCCATCTTCCCACAGGGCTAAAGAAGATAAGCTAGAGATCAAGAAGATGGCTTTCCATGCTCTG
AATGAAGTTCTAAACTTTCTGGATCCTCCTTCCTCAGAAGGGTCGAAAGAAGCGGCGGAGAAGCCT
CGAAGGGGGCGTGGTCGCCCAAGAAAACACTAG

SEQ ID NO: 144, Proteína - Oryza sativa

MLVNDKVSKTEKKRKRGDTGAAENNGKGKKMLTGENALMCHQCQRNDKGRVVWCKTCNNKRFCVPC
INQWYPDLPENEFAAKCPYCRKNCNCKACLRMRGVEEQPPRKEISKENQIRYACHVLRLLRPWLIE
LRQEQMAEKELEAKIQGVSVDQIKVEQAVCDLDERVYCNRCSTSIVDFHRSCKHCFYDLCLTCCQE
LRKGEIPGGEEVEILDPEERDKDYAFGKILSDGENQRDSLKCRSDTQNSESNKGMASDENQKKALL
LWKANSNGSIPCPRKEKEDCSFSSLDLKCLFPEKLLPELEDRSEKVFWSETFAKELGRTSELCPCF
DHSGKIRSDSKKLRQAANREDSSDNYLYCPVATDIQDADLLHFQMHWAKGEPVVVSDTLKLTSGLS
WEPMVMWRAVRERTKGKAEDEQFAVRAVDCLDWCEVEINIHMFFMGYTRGRTHPRTYWPEMLKLKD
WPPSSSFDQRLPRHGAEFISALPFPEYTDPRYGPLNLAVKLPGGVLKPDLGPKTYIAYGFSEELGR
GDSVTKLHCDMSDAVNILTHTAEVPCETYDAVQIKNTQKKMKMQDDMEIYGMIESGSELKPSACPV
ELGNKAVGEAPKASCSKENVHTLKDKSNGLDINASPPDDAGGDARDEALSYESVVHSDVAQCPNHN
HETNNSDDARIGAQRCQKKAKGRPPKTGSGVSEHQESGGALWDIFRREDSEKLQDFLRKHAPEFRH
IHCNPVKQVIHPIHDQAFYLTAEHKRKLKEEYGVEPWTFEQKLGEAVLIPAGCPHQVRNLKSCIKV
ALDFVSPENVGECVRLTKEFRRLPSSHRAKEDKLEIKKMAFHALNEVLNFLDPPSSEGSKEAAEKP
RRGRGRPRKH

FIGURA 10 CONT.

SEQ ID NO: 145, DNA - *Oryza sativa*

ATGCCGCCCCAGCCTCCGCCGGCGGCGTCGGCGTCGGCCAGTGCCCCGGATCCCGCCGTCCCGGCT
TGGCTCAGGGGGCTCCCCCGCGCGCCCGAGTACCGCCCGACGGAGTCCGAGTTCGCCGACCCCATC
GCCTTCCTCTCCCGCGTGGAGCGCGAGGCCGCCGCCTACGGCATCTGCAAGGTCATCCCGCCCCAC
CCGCGCCCTTCCCGCCGCTTCGTCTTCGCCCACCTCAACCGCTCCCTCGTCTCCTCCTGCGACGCC
CCCGCGCCCTCTCCCGCCGCCGCCTCCGACTCGTCCATTCCTCCATCCTCCTCCTCGCCGCCGCCG
GTCTCCGCCGCGGTGTTCACGACACGGCACCAGGAGCTCGGCAACCCGCGGCGTGGACGCCCGACG
CCGCAGGTGCTTAAGCAGGTGTGGCAGAGTGGGGAGCGGTACACGCTCGACCAGTTTGAGTCCAAG
TCCCGCGCCTTCTCAAAGACGCACCTCGCTGGACTCCACGAACCGACTGCGCTCGCGGTGGAATCA
CTCTTCTGGAAGGCGTCGGCGGACCGTCCTATCTATATTGAGTACGCCAATGATGTCCCTGGCTCT
GGGTTTGCTGCACCCGTGCAGTTGCAACGCAAGAAGAAGCAGAAAAGAGAGACTGCCCCGATGGAC
GAATGGGAGAAGAGTTCAGGCTGGAGGCTGTCAAATAGCCCATGGAACCTGCAAGCAATTGCGCGG
GCTCCTGGTTCGCTCACACGGTTTATGCCTGACGATGTTCCTGGGGTGACTTCTCCAATGGTTTAC
ATTGGCATGCTGTTCAGCTGGTTTGCGTGGCATGTGGAGGATCATGATCTGCATAGTCTCAACTTC
CTCCACACTGGCGCACCTAAGACGTGGTATGCGGTTCCTGGTGATAGGGCTGTTGAGCTTGAGGAA
GTTATCCGTGTACATGGCTATGGAGGCAACACTGATCGGATCGCGTCACTAGCAGTGCTTGGTGAG
AAAACAACACTAATGTCCCCAGAGGTCCTTATAGACAATGATTGGTGCAATATCCTGGTGAGTTTG
TGGTGA

SEQ ID NO: 146, Proteína - *Oryza sativa*

MPPQPPPAASASASAPDPAVPAWLRGLPRAPEYRPTESEFADPIAFLSRVEREAAAYGICKVIPPH
PRPSRRFVFAHLNRSLVSSCDAPAPSPAAASDSSIPPSSSSPPPVSAAVFTTRHQELGNPRRGRPT
PQVLKQVWQSGERYTLDQFESKSRAFSKTHLAGLHEPTALAVESLFWKASADRPIYIEYANDVPGS
GFAAPVQLQRKKKQKRETAPMDEWEKSSGWRLSNSPWNLQAIARAPGSLTRFMPDDVPGVTSPMVY
IGMLFSWFAWHVEDHDLHSLNFLHTGAPKTWYAVPGDRAVELEEVIRVHGYGGNTDRIASLAVLGE
KTTLMSPEVLIDNDWCNILVSLW

SEQ ID NO: 147, DNA - *Oryza sativa*

ATGGAGGCGGCGGCGGCGGCGGCCGCCGTGCCGGAGGAGCTGCGGTGCAAGCGCTCCGACGGCAAG
CAGTGGCGATGCAGCGCGCCCTCCATGCCCGACAAGACCGTCTGCGAGAAGCACTACGTCCAGGCC
AAGAAGCGCGCCGCCTCCTCCGCGCTCCGGGCCTCCCTCCGCAGGTCCTCCGCCTCCGCCTCCGCC
GCGCGGGGCACGACCCCGCCTGCGCGGATGGCGGTCGCGAGGCCCATCTACGGCAGGGTCGCGGGG
GAGCCGGTGTACGTGGCGGAGCCGGCGCTGCCGCCGCCGCCGCCGCCGCCACGGAGGAGGCAGCCG
GTCCACGGGCTGCCCATGGGCAATGCTGCCGGCGCGCGCACCGCGGCGGAGCTGGTCGGGAGGGGC
TCGGCGGGGCTGGTTGCCTGCAGCTCGGCGGCGGGGGCGGCAGCGGCCGCAACCTGCCATCAGTGC
CGGAGGGTCGCTAATACCATTTGTTGTACAAGCTGCGACAGAAGGGGGTACTGCACCAACTGCATC
TCGAGATGGTACTCTGATATCCCAATTGATGATGTTCGAAAGGTTTGTCCAGCATGTCGTGGCATT
TGTAACTGCAGAGTTTGCTTACTAGGAGATAATGTAATAAAGGCAAGGGTTCAGGAGATATCTGCT
GTAGATAAGTTGGAATATCTTCATAGCATTCTGGCATCTGTCCTTCCAGTATTAAAGCAGATTTAT
TCTGATCAATGTTTCGAAATTGGTGTTGATACAAAAGCTTATGGACTTAGGACAGATATAATCAGG
GCGAAGGTCAATCCTGATGAGCAAATGTGCTGTGACTTCTGCAAAGTGCCAGTATTTGATTATCAC
CGACACTGCCCAAGGTGTTTGTATGACTTATGTCTTGACTGTTGTCGAGACATAAGGCGATCTAGG
ACCAGTGTTGCAAGAGGAGAATATGCTGAAGGTCGTGTGGTAGATAGAAGTAAAGATACTTCGAAT
AAAAGAGCAAGAATGGAACCATCTGCAGAAAGTGCAAATGATAAGTCAGTCCCACAACGAAGGGAC
ATAAAAAATATTGACATTAGATCTTTATTCCCTACATGGAGAGTCAATAATGATGGAAGCATTACT
TGTGGGCCTCATGAGGCTGGTGGTTGTGGCTCCTCGAAGTTAGTGTTAAGGCGAATATTCAAAATA
AACTGGATTAGTAAGCTTGTAAAAAATTCTGAGGAAATGGTCAATGGTTGTAAAGTACATGTTCTT
GAGAATGGATGCTCATCCTGCAATGATGGCAGAACATTAGAGTTAACTGGTCATCGTAACTTTGGT

FIGURA 10 CONT.

GTCTCAACATGTTCAAATAATGGTGGTATTGATCGGTTCTGTGTGTTCTCTCCTGTATTAGAGGAC
TTGAAATCTGAAGGTATTATTCATTTTCGTAAACACTGGATAAAAGGAGAACCTGTAGTTATCAGG
AATGCGTTTGAGCCTTCTCTGTCATCAAGCTGGGACCCACTAAATATTTGGCGAGGAATCCAGGAA
ATCATGGATGAAGAAGTGGATGACGATGTCATTGTCAAAGCTGTGGATTGCTCAAACCAAGCAGAG
GTGGATATTGAGCTGAAACAGTTTATCAAAGGTTATTCAGATGGTCACAAGGGGGAAGATGGGGAA
TTGATGATGCTGAAATTGAAAGAGTGGCCCCCACCCAGTGTATTGGAGGAGTTTCTGCTATGCCAA
AGACCAGAATTTATTGTCAATTTTCCATTAGTTGATTTTATACATTCCAGGTGGGGGCTCCTAAAT
CTTTCTGCTAAATTGCCCCCAGACACCCTACAACCTGAAGTTGGCTTGAAGCTGTTAATTGCATAT
GGAAGACATCAAGAAGCTGGTAAAGGTGATTCAGTGACAAATCTAATGATTAACATGGCTGATGTG
GTGCATATGCTAATGCATACAGCCAAAGGGCATGACGTATGTCCAAAGAGGCTACAACCTGAGCGA
TCTGAAAAGATTGCCAATGGAATGACAATGCATGTAAATGCTCATGCACCTGTTCAAAATTTGAAT
GTGGATATGGGGGAACAATCACCTGACCATGTAAGCTCAAAGTTTGATGAAAGAGCGCATGCGTCT
GCCTTGCGATTACAAGAGAAGTCTTCAGATGCCAAACTTAATTGTGGTTTCGAAGGCTCTTCAACT
GAGTTATCTTGCTCGTCACATTCAGAGGAACCAAAAGTTAATGGTTCAGAAAGAAGTCAGGCTGGT
TCTGTTTGGGATGTGTTCCGCAGGCAGGATATTTCAAAGCTGAATGAATATTTAACTGCTAACTGG
GAAGAACTGGCAGCTAGTAGTCAGGTTAAGAATCCTATTTATGAACAATCTATATATCTCAACAAG
TATCATAAAAGGATACTGAAGGATCAATATGGAATTGAACCCTGGACATTCCAACAACATATCGGT
GAGGCTGTATTTGTTCCTGCTGGCTGCCCATTCCAAGTGAAAAATCTCCAGTCTACGGTTCAATTG
GCTCTTGATTTCCTGTCACCGGAAAGTTTGGGGGAGTCGGCCCGAATGGCCCAGGAGATTCGCTGC
CTACCAAATGATCATGATGCAAAACTGAAGATGCTCGAGATTGGAAAAATCTCTTTATATGCAGCT
AGTTCTGCTGTCAGAGAAATTCAGAGAATAACCCTCGATCCCAAGTTTAATCTAGACCTTAAATTC
GAGGATCAGAATCTAACTCAGGCGGTTTCTGAGAACTTGGCTAGAGTCACCAAACAACGGAATGTA
CCATGCAGCTGA

SEQ ID NO: 148, Proteína - *Oryza sativa*

MEAAAAAAAVPEELRCKRSDGKQWRCSAPSMPDKTVCEKHYVQAKKRAASSALRASLRRSSASASA
ARGTTPPARMAVARPIYGRVAGEPVYVAEPALPPPPPPPRRRQPVHGLPMGNAAGARTAAELVGRG
SAGLVACSSAAGAAAAATCHQCRRVANTICCTSCDRRGYCTNCISRWYSDIPIDDVRKVCPACRGI
CNCRVCLLGDNVIKARVQEISAVDKLEYLHSILASVLPVLKQIYSDQCFEIGVDTKAYGLRTDIIR
AKVNPDEQMCCDFCKVPVFDYHRHCPRCLYDLCLDCCRDIRRSRTSVARGEYAEGRVVDRSKDTSN
KRARMEPSAESANDKSVPQRRDIKNIDIRSLFPTWRVNNDGSITCGPHEAGGCGSSKLVLRRIFKI
NWISKLVKNSEEMVNGCKVHVLENGCSSCNDGRTLELTGHRNFGVSTCSNNGGIDRFCVFSPVLED
LKSEGIIHFRKHWIKGEPVVIRNAFEPSLSSSWDPLNIWRGIQEIMDEEVDDDVIVKAVDCSNQAE
VDIELKQFIKGYSDGHKGEDGELMMLKLKEWPPPSVLEEFLLCQRPEFIVNFPLVDFIHSRWGLLN
LSAKLPPDTLQPEVGLKLLIAYGRHQEAGKGDSVTNLMINMADVVHMLMHTAKGHDVCPKRLQPER
SEKIANGMTMHVNAHAPVQNLNVDMGEQSPDHVSSKFDERAHASALRLQEKSSDAKLNCGFEGSST
ELSCSSHSEEPKVNGSERSQAGSVWDVFRRQDISKLNEYLTANWEELAASSQVKNPIYEQSIYLNK
YHKRILKDQYGIEPWTFQQHIGEAVFVPAGCPFQVKNLQSTVQLALDFLSPESLGESARMAQEIRC
LPNDHDAKLKMLEIGKISLYAASSAVREIQRITLDPKFNLDLKFEDQNLTQAVSENLARVTKQRNV
PCS

SEQ ID NO: 149, DNA - *Oryza sativa*

ATGGCGGCGGGGAACGGCCGGATGGAGGCGGCGCTGGGCTGCCTCGCGGCGCTCCCCGACGAGGTG
CTCTGCGCCGTCGTCGACCTCCTCCCGCCCACCGACGTCGGCCGCCTCGCCTGCGTCAGCAGTGTC
ATGTACATACTTTGCAATGAGGAGCCTCTCTGGATGAGCAAGTGTCTTTCAGTTGGGGGTCTTCTT
GTGTATAGAGGTTCTTGGAAGAAAACAGCATTGTCTAGACTTAATCTTTGTTCAGAAAATGATGAG
ATTTACCAGAAGCCTCGCCATTTTGATGGGTTCAATTCCATGCACTTATACAGGAGATGGTACAGA
TGTTTTACTAATTTGAGTAGCTTTTCCTTTGATAATGGGCACGTTGAAAGGAAAGATGACCTTTCT

FIGURA 10 CONT.

CTAGACCAATTTCGCGCTCAGTATGATGGAAAATGTCCAGTTTTGCTTACTAAACTGGCTGAAACC
TGGCCAGCAAGGACTAAATGGACAGCGCAGCAACTGACACATGATTATGGTGAAGTTCCCTTTAGG
ATATCTCAGAGAAGCCCTCAAAAGATAAAAATGAAACTAAAAGATTATGTTTTTTACATGGAACTC
CAGCATGACGAAGATCCACTTTACATATTTGATGATAAGTTTGGAGAATCAGCACCTACACTATTG
GAAGATTACAGTGTCCCTCATCTATTTCAAGAAGATTTCTTTGAAATCATGGATTACGACCAACGA
CCAGCTTTCAGATGGCTTATTATTGGACCAGAGAGATCAGGTGCTTCTTGGCATGTTGATCCAGGG
TTGACCAGTGCCTGGAATACTCTTCTTTGTGGCCGAAAAAGGTGGGCAATGTACCCTCCTGGAAGA
GTACCAGGTGGTGTCACAGTACATGTCAGTGATGAAGATGGTGATGTTGACATTGAAACTCCTACA
TCTTTGCAGTGGTGGCTAGATATCTACCCAAATCTTGCTGAGCATGAGAAACCACTGGAATGCACA
CAATTACCAGGAGAGACCATATTTGTTCCTAGTGGGTGGTGGCATTGTGTTTTGAACCTTGACATG
ACAATTGCTGTCACGCAAAATTTTGTCAACCAATCAAATTTTAAGCATGTATGTTTGGACATGGCA
CCTGGTTACTGTCACAAAGGAGTTTGCCGTGCTGGCTTACTTGCTGCTCCAGACAAATCTATTAGA
GATATTGAAAATCTTCCTAGTATAACGAGTAGATTGAACCACTCTGACATGGCCTGTAAGGAAAAA
AGACTGAAAAGTTCAGAGCCTATAAGAACTTCAAATAATGCAAATCAGTGTTCTGCATTTGAGTTC
TCAGATGTTCATGAAAACTTGGGGGACCAAGTTTTTTCGTATGATATAGATTTCTTATCCCAATTC
CTTGAGAAAGAAAAGGATCACTATTCTTCTGTCTGGAGCCCTACTAATTCAATTGGCCAGAGAGAA
GCAAGAGAATGGCTACGTAGGCTATGGGTTCTTAAACCTGAATTGAGAGAACTAATATGGAAGGGT
GCATGTCTAGCAATTAATGTAGACAAGTGGTATTCATGCTTAGAGGAAATAAGTGCATGCCATAGT
TTACCACCACCTTCTGAAGATGAGAAGCTTCCTGTTGGCACAGGTAGCAACCCAGTCTTCATTGTT
TCTGGCAATGTGATCAAAATTTATGCTGAAGGAGGGTTGGGTTATTCTATACATGGTTTGGGCACA
GAGCTTGAGTTCTATGATCTTCTGCAAAAACTTGGCTCGCCATTGATCAACCATGTCCCTGAGATC
ATTGCAAGTGGCTTTCTTGTGTACCTGGATGGTGTCTACAAGACAGTTCCATGGGATGGAAACGGA
ATACCAGATGTTCTAGCTAAATACTACTCTTTGGAGGTGTCTTATGCAAACGGCTCTTTTCCTCTT
GGATTATGGAGCAAGCAACTGTTTGGATTGAGTAATTCAACTGATGCTCCAGACAGACCAATTTGT
CCTTACATGGTTACCAGAAAATGCAAAGGGGATATTTTTGCTCGCATACGTGATAAATTGACCAAG
ACTGATGTTTTGAATCTTGCATCATCCTTGGGAGTTCAAATGCGAAATATTCATCAATTACCCCTT
CCACATGTGGAACACATATCCAAATCTGGGAACGAAGATATCAAAGCAAAGGAAAATTCAATTTCT
GATGTCACTCATGTTCCGCCTGAATGGAAACAAGTAGTTTCTACTCTAGACAGGAGAAAGAAAAGT
ATAAAGAAGCATCTAAGTAACTGGGGTGGTTCAATTCCACAGGTTCTAATTGAGAAGGCTGAAGAA
TATCTCCCTGACGACATCCGCTTTCTTATCAAGTTTGTTAAGGACGATGATGGTGATTCAGTCTAT
GTGGTACCTTCTTGGATACATTCAGATATAATGGATGATAACATTCTCATTGAGGGGACCACAGAA
CCAGGAACTTCCACTGATTGCATTGCCGTTGAAGATCTGAACAAAATGGATGCAATTCATATCATT
GATTTCAGTGATCTGTCCATTGGGGATCCTCTATGTGACTTAATTCCACTGCACTTGGATGTATTC
CGTGGTGATATTGATCTTCTCAGGCAGTTTTTACGAAGCTATCAGCTTCCTTTTCTGAGAGCAGAA
TCAAATAAAGATATATACAAGTCAATACAAAATTCTAAATTCAGCAGGGCATCGTATCGTGCGATG
TGCTACTGCATACTTCACGAGGACAACGTCCTGGGAGCCATATTTAGCCTGTGGAAGGATCTGGGC
ACCGCGACGTCATGGGAAGATGTTGAACACTTGGTTTGGGGAGAGCTGAATCAATACCAGCAGTCA
TGCAGCGTGGGCGAAATTAACTGA

SEQ ID NO: 150, Proteína - *Oryza sativa*

MAAGNGRMEAALGCLAALPDEVLCAVVDLLPPTDVGRLACVSSVMYILCNEEPLWMSKCLSVGGLL
VYRGSWKKTALSRLNLCSENDEIYQKPRHFDGFNSMHLYRRWYRCFTNLSSFSFDNGHVERKDDLS
LDQFRAQYDGKCPVLLTKLAETWPARTKWTAQQLTHDYGEVPFRISQRSPQKIKMKLKDYVFYMEL
QHDEDPLYIFDDKFGESAPTLLEDYSVPHLFQEDFFEIMDYDQRPAFRWLIIGPERSGASWHVDPG
LTSAWNTLLCGRKRWAMYPPGRVPGGVTVHVSDEDGDVDIETPTSLQWWLDIYPNLAEHEKPLECT
QLPGETIFVPSGWWHCVLNLDMTIAVTQNFVNQSNFKHVCLDMAPGYCHKGVCRAGLLAAPDKSIR
DIENLPSITSRLNHSDMACKEKRLKSSEPIRTSNNANQCSAFEFSDVHENLGDQVFSYDIDFLSQF
LEKEKDHYSSVWSPTNSIGQREAREWLRRLWVLKPELRELIWKGACLAINVDKWYSCLEEISACHS

FIGURA 10 CONT.

LPPPSEDEKLPVGTGSNPVFIVSGNVIKIYAEGGLGYSIHGLGTELEFYDLLQKLGSPLINHVPEI
IASGFLVYLDGVYKTVPWDGNGIPDVLAKYYSLEVSYANGSFPLGLWSKQLFGLSNSTDAPDRPIC
PYMVTRKCKGDIFARIRDKLTKTDVLNLASSLGVQMRNIHQLPLPHVEHISKSGNEDIKAKENSIS
DVTHVPPEWKQVVSTLDRRKKSIKKHLSNWGGSIPQVLIEKAEEYLPDDIRFLIKFVKDDDGDSVY
VVPSWIHSDIMDDNILIEGTTEPGTSTDCIAVEDLNKMDAIHIIDFSDLSIGDPLCDLIPLHLDVF
RGDIDLLRQFLRSYQLPFLRAESNKDIYKSIQNSKFSRASYRAMCYCILHEDNVLGAIFSLWKDLG
TATSWEDVEHLVWGELNQYQQSCSVGEIN

SEQ ID NO: 151, DNA - *Oryza sativa*

ATGGAGATGGAGGAGGCGGTAGACGGCAAGCACCCGAAGCGTGGGAGGGGCAGGCCGAGGGGACGA
CGGGGCAGGGGCAGGGGCAGGGGGAGGGGCGGCCGCTCCCTCGCGTCGCCGGCGGCGGGGCCCGGA
GACCAAGGGCCGCGGCGGCGGCGCGGGGTCGTCCCGGCGGCGGCGGCGGCGGCGGGGGGAAGGGCG
CTGAGGGAGAGGAGGCCGGCGCCTGGTGCCTACCGCGAGAGCGGAGCTGATAATGACGATGATGGC
GGCGGCGACGATGAGCACGACGAGCAGAATGATGATGGTGCTGAAAAATCGGATAATCAAGTTGTG
GATTCTCTTAATGAACCCAACAGAAGTAATACAGGGAAGAAGCGGGGAAGACCAAAGAAAGTGAAA
GCAGAGCAGGAGGACAGTAACCAACTCTCCAACGGGAAACACCTTGGCGAAAACAATGGGAATGAT
GAAGCGATAATGATGAAGCCTTCAAAAGAATCGAAGAAAAGAGGTGCAGGAAAGAAACAAGAGGAA
GAAGAGAATAATACCATATCCATTGAGGATGAAATGTGTGATGCGAACAATAAGAAGGGGAAGAAG
ATGCTCACTGGGGAAAATGCTTTGATGTGCCACCAGTGTCAAAGAAATGACAAAGGGAGGGTGATT
TGGTGTAAATCATGCAACAATAAGCGCTTCTGTGAGCCATGCATGAAACGATGGTACCCAGGTTTG
TCAGAAGTTGATTTTGCTGCAAAATGCCCATATTGTAGAAAGAACTGTAATTGCAAGGCATGCCTA
AGAATGATAGGAGTTGAAAAGCCCCCTGAAAAGAAGATTTCTGAGGAAAATCAACGACGTTATGCT
TTTCGTATTGTAGACTTGTTGCTTCCTTGGTTGAAAGAACTTCAACAGGAGCAGATGAAGGAGAAG
GAATTAGAGGGTAGATTACAAGGTGTTTCTATGGATGAAGTAAAGTTGGAACAAGCTGATTGTGAT
ATGGATGAGCGTGTTTATTGTGATAGGTGCAAAACTTCGATAGTTGATTTTCATAGAAGCTGCAAG
GCTTGTTCGTATGACTTGTGCCTTGCATGCTGCTGGGAGCTTCGCAAAGGTGAGATTCCAGGAGGA
GAAGAGGCAAAGAGTGTGCAGTGGGAAGAGAGAGGTCAGAAGTATGTTTTTGGTAATATTTCCAAG
GATGAGAAGAAAAGGGTATCTTCGAAGAGGCACATGGAGACCCCCAGTACTGAAACTTGCAATGAC
ATGGCTGTTGCTGGGGACCCAAATAACCCTTTGTTACTGTGGAAAGCTAATAGTGATGGCAGTATA
CCTTGTCCACCAAAGGAAATAGGGGGCTGCGGTGCTTCATCTTTGGTACTCAGATGCTTATTACCA
GAAATTATGCTTTCTGAGTTAGAACACAGGGCTAACAAAGTTATTAAGAGAGAAGCATTTGATAAA
GCAATAAACGAAACAAGTGATCAGTGCCCTTGCTTTTATCACACAAGCAAGATAAGAACAAATGCT
ACTCGAGAGGCAGCAAACAGAAAGGGCTCAAGTGATAACTACTTGTACTGTCCAGATGCCAATAAT
ATCCAGGAGGATGACCTGTCGCACTTTCAGATGCATTGGTCAAAAGGTGAACCGGTTATTGTTTCT
GATGCTCTTCGGTTAACATCTGGTCTGAGCTGGGAGCCATTGGTTATGTGGCGGGCATTGCGAGAG
AAGAAAACCAATGGTGATGTTGAAGATGAGCACTTTGCTGTTAAGGCAGTGGATTGCCTTGATTGG
AATGAGGTGGAAATTAACATACACATGTTCTTTATGGGGTATATGAGAGGTAGAAGACATCCGATG
ACCTTTTGGCCTGAGATGCTTAAGCTAAAAGATTGGCCACCATCTAGTATGTTTGACCAGAGGTTA
CCTCGCCATGGTGCTGAGTTTATAACTGCATTGCCATTTCCCGAGTATACTGATCCACGATATGGC
CCGCTAAATCTTGCTGTCAGGCTTCCTGCTGGTGTACTGAAGCCTGATCTTGGGCCAAAAACTTAT
ATTGCTTATGGATGTTATGAAGAGTTAGGCAGGGGTGACTCTGTGACCAAGCTTCATTGCGACATG
TCTGATGCGGTAAATATCTTGATGCACACAGCTGAAGTGTCCTATGACACTGAACAGCTTGACAAG
ATAGCAAAAATTAAAATGAAAATGAGAGAACAAGATCTTCATGAACTGTTTGGGGTTTCAGAATCA
GGCGCCAAGGGTAAAGCTGATGATGAAGCATCGAAAATCTCATGTAACATGGAAAACAAACACACC
TCTAACCAAAGTACTAAGGGTTTGGACATTAATGCTTTACCACCTGATGATTCTGGAAGTGATATT
GGGGATAAACCATCATTTTGTCAATCTGAGGTAGAAAGTGAATTAACACAATGTTCAAAACACAAT
CACGAGGTTAATAGTTCTGTCAAGATGCATGCTGGAGCTCATTGTACTTCAGACAACCAAGGATAC
ATTGACAGGAGTGGATTCAAACGCAAAGATTCAGACTGTTCAGACCAACAGAAAACTGGTGGCGCT

FIGURA 10 CONT.

TTGTGGGATATTTTCCGAAGAGAAGATTCTGAGAAGTTACAAGATTATCTTCGGAAGCATGCCTCA
GAATTTCGGCACATACACTGCAATCCAGTGAAAAATGTTTCTCACCCAATTCATGACCAGACTTTC
TATTTAACTGTGGAGCATAAGAGAAAGCTCAAGGAAGAACACGGTGTTGAACCATGGACATTCGAG
CAGAAGCTAGGCGATGCAGTTTTCATTCCTGCTGGATGTCCACACCAAGTGAGAAATTTAAAGTCT
TGCATCAAGGTTGCTCTAGACTTTGTTTCCCCTGAAAATGTTGGTGAGTGTGTTAAGCTGACCGGA
GAGTTTAGACGTCTTCCATCTGATCACAGGGCTAAAGAAGATAAACTAGAGATTAAGAAGATTGCT
CTCAATGCTCTTAAAGAAGTCGTAAATTTCTTAGATCCTTTACCAAAAGGGTCAAAGAACAGGGAT
GAAGTGGTAGAAGTGACCAAACCAAAAAGGAAATACGGTAACCGAAGAGGTGACCTGAAGAGTGGG
GAAGACCAACCCATTGATGAATCCATAGAAGAAAGGAAGCCTAAAAAGCGAGGGAGATCTAAACGG
TGA

SEQ ID NO: 152, Proteína - *Oryza sativa*

MEMEEAVDGKHPKRGRGRPRGRRGRGRGRGRGGRSLASPAAGPGDQGPRRRRGVVPAAAAAAGGRA
LRERRPAPGAYRESGADNDDDGGGDDEHDEQNDDGAEKSDNQVVDSLNEPNRSNTGKKRGRPKKVK
AEQEDSNQLSNGKHLGENNGNDEAIMMKPSKESKKRGAGKKQEEEENNTISIEDEMCDANNKKGKK
MLTGENALMCHQCQRNDKGRVIWCKSCNNKRFCEPCMKRWYPGLSEVDFAAKCPYCRKNCNCKACL
RMIGVEKPPEKKISEENQRRYAFRIVDLLLPWLKELQQEQMKEKELEGRLQGVSMDEVKLEQADCD
MDERVYCDRCKTSIVDFHRSCKACSYDLCLACCWELRKGEIPGGEEAKSVQWEERGQKYVFGNISK
DEKKRVSSKRHMETPSTETCNDMAVAGDPNNPLLLWKANSDGSIPCPPKEIGGCGASSLVLRCLLP
EIMLSELEHRANKVIKREAFDKAINETSDQCPCFYHTSKIRTNATREAANRKGSSDNYLYCPDANN
IQEDDLSHFQMHWSKGEPVIVSDALRLTSGLSWEPLVMWRALREKKTNGDVEDEHFAVKAVDCLDW
NEVEINIHMFFMGYMRGRRHPMTFWPEMLKLKDWPPSSMFDQRLPRHGAEFITALPFPEYTDPRYG
PLNLAVRLPAGVLKPDLGPKTYIAYGCYEELGRGDSVTKLHCDMSDAVNILMHTAEVSYDTEQLDK
IAKIKMKMREQDLHELFGVSESGAKGKADDEASKISCNMENKHTSNQSTKGLDINALPPDDSGSDI
GDKPSFCQSEVESELTQCSKHNHEVNSSVKMHAGAHCTSDNQGYIDRSGFKRKDSDCSDQQKTGGA
LWDIFRREDSEKLQDYLRKHASEFRHIHCNPVKNVSHPIHDQTFYLTVEHKRKLKEEHGVEPWTFE
QKLGDAVFIPAGCPHQVRNLKSCIKVALDFVSPENVGECVKLTGEFRRLPSDHRAKEDKLEIKKIA
LNALKEVVNFLDPLPKGSKNRDEVVEVTKPKRKYGNRRGDLKSGEDQPIDESIEERKPKKRGRSKR

SEQ ID NO: 153, DNA - *Oryza sativa*

ATGGAGATGGAGGAGGCGGTAGACGGCAAGCACCCGAAGCGTGGGAGGGGCAGGCCGAGGGGACGA
CGGGGCAGGGGCAGGGGCAGGGGGAGGGGCGGCCGCTCCCTCGCGTCGCCGGCGGCGGGGCCCGGA
GACCAAGGGCCGCGGCGGCGGCGCGGGGTCGTCCCGGCGGCGGCGGCGGCGGCGGGGGGAAGGGCG
CTGAGGGAGAGGAGGCCGGCGCCTGGTGCCTACCGCGAGAGCGGAGCTGATAATGACGATGATGGC
GGCGGCGACGATGAGCACGACGAGCAGAATGATGATGGTGCTGAAAAATCGGATAATCAAGTTGTG
GATTCTCTTAATGAACCCAACAGAAGTAATACAGGGAAGAAGCGGGGAAGACCAAAGAAAGTGAAA
GCAGAGCAGGAGGACAGTAACCAACTCTCCAACGGGAAACACCTTGGCGAAAACAATGGGAATGAT
GAAGCGATAATGATGAAGCCTTCAAAAGAATCGAAGAAAAGAGGTGCAGGAAAGAAACAAGAGGAA
GAAGAGAATAATACCATATCCATTGAGGATGAAATGTGTGATGCGAACAATAAGAAGGGGAAGAAG
ATGCTCACTGGGGAAAATGCTTTGATGTGCCACCAGTGTCAAAGAAATGACAAAGGGAGGGTGATT
TGGTGTAAATCATGCAACAATAAGCGCTTCTGTGAGCCATGCATGAAACGATGGTACCCAGGTTTG
TCAGAAGTTGATTTTGCTGCAAAATGCCCATATTGTAGAAAGAACTGTAATTGCAAGGCATGCCTA
AGAATGATAGGAGTTGAAAAGCCCCCTGAAAAGAAGATTTCTGAGGAAAATCAACGACGTTATGCT
TTTCGTATTGTAGACTTGTTGCTTCCTTGGTTGAAAGAACTTCAACAGGAGCAGATGAAGGAGAAG
GAATTAGAGGGTAGATTACAAGGTGTTTCTATGGATGAAGTAAAGTTGGAACAAGCTGATTGTGAT
ATGGATGAGCGTGTTTATTGTGATAGGTGCAAAACTTCGATAGTTGATTTTCATAGAAGCTGCAAG
GCTTGTTCGTATGACTTGTGCCTTGCATGCTGCTGGGAGCTTCGCAAAGGTGAGATTCCAGGAGGA
GAAGAGGCAAAGAGTGTGCAGTGGGAAGAGAGAGGTCAGAAGTATGTTTTTGGTAATATTTCCAAG

FIGURA 10 CONT.

GATGAGAAGAAAAGGGTATCTTCGAAGAGGCACATGGAGACCCCCAGTACTGAAACTTGCAATGAC
ATGGCTGTTGCTGGGGACCCAAATAACCCTTTGTTACTGTGGAAAGCTAATAGTGATGGCAGTATA
CCTTGTCCACCAAAGGAAATAGGGGGCTGCGGTGCTTCATCTTTGGTACTCAGATGCTTATTACCA
GAAATTATGCTTTCTGAGTTAGAACACAGGGCTAACAAAGTTATTAAGAGAGAAGCATTTGATAAA
GCAATAAACGAAACAAGTGATCAGTGCCCTTGCTTTTATCACACAAGCAAGATAAGAACAAATGCT
ACTCGAGAGGCAGCAAACAGAAAGGGCTCAAGTGATAACTACTTGTACTGTCCAGATGCCAATAAT
ATCCAGGAGGATGACCTGTCGCACTTTCAGATGCATTGGTCAAAAGGTGAACCGGTTATTGTTTCT
GATGCTCTTCGGTTAACATCTGGTCTGAGCTGGGAGCCATTGGTTATGTGGCGGGCATTGCGAGAG
AAGAAAACCAATGGTGATGTTGAAGATGAGCACTTTGCTGTTAAGGCAGTGGATTGCCTTGATTGG
AATGAGGTGGAAATTAACATACACATGTTCTTTATGGGGTATATGAGAGGTAGAAGACATCCGATG
ACCTTTTGGCCTGAGATGCTTAAGCTAAAAGATTGGCCACCATCTAGTATGTTTGACCAGAGGTTA
CCTCGCCATGGTGCTGAGTTTATAACTGCATTGCCATTTCCCGAGTATACTGATCCACGATATGGC
CCGCTAAATCTTGCTGTCAGGCTTCCTGCTGGTGTACTGAAGCCTGATCTTGGGCCAAAAACTTAT
ATTGCTTATGGATGTTATGAAGAGTTAGGCAGGGGTGACTCTGTGACCAAGCTTCATTGCGACATG
TCTGATGCGGTAAATATCTTGATGCACACAGCTGAAGTGTCCTATGACACTGAACAGCTTGACAAG
ATAGCAAAAATTAAAATGAAAATGAGAGAACAAGATCTTCATGAACTGTTTGGGGTTTCAGAATCA
GGCGCCAAGGGTAAAGCTGATGATGAAGCATCGAAAATCTCATGTAACATGGAAAACAAACACACC
TCTAACCAAAGTACTAAGGGTTTGGACATTAATGCTTTACCACCTGATGATTCTGGAAGTGATATT
GGGGATAAACCATCATTTTGTCAATCTGAGGTAGAAAGTGAATTAACACAATGTTCAAAACACAAT
CACGAGGTTAATAGTTCTGTCAAGATGCATGCTGGAGCTCATTGTACTTCAGACAACCAAGGATAC
ATTGACAGGAGTGGATTCAAACGCAAAGATTCAGACTGTTCAGACCAACAGAAAACTGGTGGCGCT
TTGTGGGATATTTTCCGAAGAGAAGATTCTGAGAAGTTACAAGATTATCTTCGGAAGCATGCCTCA
GAATTTCGGCACATACACTGCAATCCAGTGAAAAATGTTTCTCACCCAATTCATGACCAGACTTTC
TATTTAACTGTGGAGCATAAGAGAAAGCTCAAGGAAGAACACGGTGTTGAACCATGGACATTCGAG
CAGAAGCTAGGCGATGCAGTTTTCATTCCTGCTGGATGTCCACACCAAGTGAGAAATTTAAAGTCT
TGCATCAAGGTTGCTCTAGACTTTGTTTCCCCTGAAAATGTTGGTGAGTGTGTTAAGCTGACCGGA
GAGTTTAGACGTCTTCCATCTGATCACAGGGCTAAAGAAGATAAACTAGAGATTAAGAAGATTGCT
CTCAATGCTCTTAAAGAAGTCGTAAATTTCTTAGATCCTTTACCAAAAGGGGATGAAGTGGTAGAA
GTGACCAAACCAAAAAGGAAATACGGTAACCGAAGAGGTGACCTGAAGAGTGGGGAAGACCAACCC
ATTGATGAATCCATAGAAGAAAGGAAGCCTAAAAAGCGAGGGAGATCTAAACGGTGA

SEQ ID NO: 154, Proteína - *Oryza sativa*

MEMEEAVDGKHPKRGRGRPRGRRGRGRGRGRGGRSLASPAAGPGDQGPRRRRGVVPAAAAAAGGRA
LRERRPAPGAYRESGADNDDDGGGDDEHDEQNDDGAEKSDNQVVDSLNEPNRSNTGKKRGRPKKVK
AEQEDSNQLSNGKHLGENNGNDEAIMMKPSKESKKRGAGKKQEEEENNTISIEDEMCDANNKKGKK
MLTGENALMCHQCQRNDKGRVIWCKSCNNKRFCEPCMKRWYPGLSEVDFAAKCPYCRKNCNCKACL
RMIGVEKPPEKKISEENQRRYAFRIVDLLLPWLKELQQEQMKEKELEGRLQGVSMDEVKLEQADCD
MDERVYCDRCKTSIVDFHRSCKACSYDLCLACCWELRKGEIPGGEEAKSVQWEERGQKYVFGNISK
DEKKRVSSKRHMETPSTETCNDMAVAGDPNNPLLLWKANSDGSIPCPPKEIGGCGASSLVLRCLLP
EIMLSELEHRANKVIKREAFDKAINETSDQCPCFYHTSKIRTNATREAANRKGSSDNYLYCPDANN
IQEDDLSHFQMHWSKGEPVIVSDALRLTSGLSWEPLVMWRALREKKTNGDVEDEHFAVKAVDCLDW
NEVEINIHMFFMGYMRGRRHPMTFWPEMLKLKDWPPSSMFDQRLPRHGAEFITALPFPEYTDPRYG
PLNLAVRLPAGVLKPDLGPKTYIAYGCYEELGRGDSVTKLHCDMSDAVNILMHTAEVSYDTEQLDK
IAKIKMKMREQDLHELFGVSESGAKGKADDEASKISCNMENKHTSNQSTKGLDINALPPDDSGSDI
GDKPSFCQSEVESELTQCSKHNHEVNSSVKMHAGAHCTSDNQGYIDRSGFKRKDSDCSDQQKTGGA
LWDIFRREDSEKLQDYLRKHASEFRHIHCNPVKNVSHPIHDQTFYLTVEHKRKLKEEHGVEPWTFE
QKLGDAVFIPAGCPHQVRNLKSCIKVALDFVSPENVGECVKLTGEFRRLPSDHRAKEDKLEIKKIA
LNALKEVVNFLDPLPKGDEVVEVTKPKRKYGNRRGDLKSGEDQPIDESIEERKPKKRGRSKR

FIGURA 10 CONT.

SEQ ID NO: 155, DNA - *Oryza sativa*

ATGATGGGGGTTACCACCACGCTCAACGAGGACACTGAACCCTCTATTCCACCTGGATTTGGACCT
TTTGCTACCCTTCCGTTATGGGGAATCCACAATGATGCCAAACCTGCTGTTACTCATTCTACTCCT
GTTCAAGCATTGCAAAGCATTAGAAAAGACAGCGAAGAATGCCAACCCAGTGCGGCTGTGTCTCGG
AGTGATACACCTTGCAGCACTTCCGGAACCCAGACATGCAGAAAATCACTGCGTAACAGACCCCCA
ATAGACTATAGCCGCTTTGAACATATATCGGATGAAGATTCTGATGTCGAAATAGTGGAAAAGGAT
GTAAGTTCAACGAGACGCAGACAACAGCTACCGAAAGGAGTACTTCGAGGATGTGCAGAATGCAGT
GACTGTCAAAAGGTTATCGCAAAATGGAATCCAGCTGGTGCACGCAGGCCTGTTCTTGATGAGGCT
CCTGTTTTCTATCCAACAGAGGAGGAATTTGAAGACACTCTAAAATACATTGAGAGTATACGGCCA
ATGGCGGAACCATATGGTATTTGCCGTATTGTCCCACCATCTTCTTGGAAGCCTCCATGCCTTCTT
AAAGATAAAAGCATATGGGAAGGATCAAAATTCTCTACTCGGGTACAAAAGGTTGACAAGCTCCAA
AACCGTAAATCATCAAAAAAGGGCAGAAGAGGTGGAATGATGAAGAGGAGAAAGCTTGCAGAGTCA
GAGGAGAACAGTGCCACTGCTCACACTCAGACAGGGATGCAGCAAAGTCCAGAGAGATTTGGATTT
GAACCTGGGCCAGAGTTCACGTTACAGACATTTCAGAAATATGCAGATGATTTCAGTAAGCAGTAC
TTTAGGAAAGATACATCGATGGATTCAGTACCATCAGTGGAAGATATTGAAGGTGAGTACTGGCGC
ATCGTTGAGGTTCCCACAGAAGAGATAGAGGTGATATATGGTGCTGATCTGGAGACTGGAACTTTC
GGCAGTGGTTTTCCAAAATTATCTCCTGAGACAAAATCTGATGCTGAGGATAAATATGCACAATCT
GGTTGGAATCTAAATAACTTGCCTAGACTACAAGGTTCAGTTCTTTCTTTCGAGGGCGGTGACATT
TCTGGTGTTCTAGTGCCTTGGGTGTATGTTGGCATGTGTTTTTCATCATTCTGCTGGCATGTTGAA
GACCATCATTTATACTCACTAAACTACATGCATTGGGGTGCCCCAAAGTTGTGGTATGGAGTTCCA
GGAAAGGATGCTGTGAATTTGGAATCTGCAATGAGGAAACATCTACCTGAATTATTTGAGGAGCAA
CCTGATTTGCTACACAACCTTGTTACCCAGTTTTCACCATCGCTGCTGAAATCTGAAGGAGTACAT
GTATACCGTTGTGTTCAGCATGAGGGCGAGTTTGTCTTGACATTCCCAAGGGCGTACCATGCTGGT
TTCAATTGTGGCTTCAATTGTGCCGAAGCTGTTAATGTGGCTCCTATTGATTGGTTACCGATTGGA
CATAATGCTGTAGAGCTTTATCGTGAGCAAGCTAGGAAAATAACCATTTCTCATGATAAGTTGTTG
TTGGGGGCTGCAAGAGAAGCAATAAGAGCTCAGTGGGATATCCTATTCCTCAAGAGGAATACTGCT
GATAATATGAGGTGGAAGAGTATATGCGGAGCTGATAGCACTATATTCAAGGCTCTTAAGGCACGA
ATTGAGACAGAGTTGGTGCAAAGGAAAACTCTAGGTGTTCCAGCTCAATCAAGGAAAATGGATGCT
GAATTCGATTCCATTGATAGGGAATGTGCCTTGTGCTACTATGATTTACATCTTTCTGCTTCTGGC
TGTCCATGCTGCCCAGAGAAATATGCTTGCCTTGTACATGCAAAGCAACTTTGCTCATGTGACTGG
GACAAAAGGTTTTTCCTATTCCGCTATGATGTCAATGAGCTAAATATCTTAGCTGATGCTTTAGGG
GGGAAATTAAGTGCCATTCATAGATGGGGCGTCTCTGATCTTGGATTAAGTTTGAGTTCATGTGTC
AAACGAGAAAAGGTCCAAGATTCCAAGACTGTTCGCAGATTAACTGATGGTCCAAGAAGGTCTTAC
ATGTCACAGGCATCAGCAGTATCCTTGGTTTCTTCTTCTACTTCCAATGAACAGAAAGATGAAGGA
AATAAGATCATGAAGATAGCTAGCCCACAGACAAATAATGTGTGCCCTTCTGTCGAGCAAAGGAAA
TCAGAGAATATTTCACCATTGAAGGAGCCATGTGTAAGGAATGAGTTGTCATGTACAACAAATTCT
GATAGTAACGGATTGCAATATAATGGAGGACTTGGAGGCCATAAAGGATCTGCACCAGGCTTGCCA
GTTTCTTCTAGCCCATCATTTTCTTCCAACGTTGCAACAAGGCCCATTAGTACTTCAAGTGTATCC
ATGAAAATTGTGCAAGGCTTGGTGGCATCTAAAAGTTGTATACAAGCTTCCTCTCGAACTGGAGAC
AGTAGATCATTGCTTGGTGAGCATCATAACAGATCACCGGCAATGATTCATGATGGAACCAACATG
AAGTCCAGTTTGGAAAGCTCAAACAATTCTTGCAGGTTGATTGCATCTGACTATAATGCAACTCCG
TGTCATTCATCCAAGGATCAGGTATTAGTAACACCAGGGACTAATGCCTCAGTAGTGACTCTGAAA
GATAGCAGCCAGGTCCATAGTGCGTCAAGTCAGCAGTTTGTCAGAACTGGCCCATGGACACAAAGT
GCTTCTCATGAAGCATCATCACCTAGTACCTCTGCTTTGAAGCCTTCTTTAGATCCCCCTGCCATG
AAAAATCTGTATGGGGGTTTTACTCAAGGCAGTGCCCATCCTGGACCTCCAAGTTTCAGTAATCAG
CAACCAAATGATGGGCGTCTTCAAAGAACATCTGAATCTCTACCAGGTGTGGAAGCTAGAGCTAGG
GGACATCCAACTGTCACGGCACAGCCTGCACTAGAAATTCACAGCAGGAATGGAGGTGCACAGAAG
GGTCCTCGCATAGCCAATGTTGTGCATCGTTTCAAGTGCTCTGTTGAACCTCTCGAAATTGGTGTT

FIGURA 10 CONT.

GTGCTATCAGGGAGGCTGTGGTCTTCAAGCCAAGCAATCTTCCCGAAAGGGTTTAGAAGCAGAGTG
AAATACTTCAGCATTGTGGATCCAATCCAAATGGCATACTACATATCGGAAATACTGGATGCTGGG
ATGCAGGGGCCTCTGTTTATGGTAAAATTAGAGAACTGTCCAGGTGAAGTTTTCATTAACTTATCT
CCAACCAAGTGTTGGAACATGGTCCGTGAAAGGCTGAACATGGAAATAAGGAGGCAACTTAATATG
GGAAAATCAAATCTTCCTACATTGCAGCCTCCAGGATCAGTTGATGGTCTTGAAATGTTTGGTTTA
TTATCACCACCAATAGTTCAGGCAATTTGGGCGCGGGACAGAGATCACATCTGTACAGAGTACTGG
AGATCAAGGCCCCATGTTCTCATTGAGGATCCAAACAATCGGCATATGTTATCTCAGGGTCCACCT
CTCCTTGCCCTGAGGGGTCTCATCCAAAGGGCTAACCGGGATGAATTGCAAGTCCTGCGGAGTTTG
ATGACGAACAGCAACAATTTGGATGATAGCTCCAGGCAACAGGCCGCGCACATTATCGAAGAGGAG
ATTGCGAAGCAATTGTGCTGA

SEQ ID NO: 156, Proteína - *Oryza sativa*

MMGVTTTLNEDTEPSIPPGFGPFATLPLWGIHNDAKPAVTHSTPVQALQSIRKDSEECQPSAAVSR
SDTPCSTSGTQTCRKSLRNRPPIDYSRFEHISDEDSDVEIVEKDVSSTRRRQQLPKGVLRGCAECS
DCQKVIAKWNPAGARRPVLDEAPVFYPTEEEFEDTLKYIESIRPMAEPYGICRIVPPSSWKPPCLL
KDKSIWEGSKFSTRVQKVDKLQNRKSSKKGRRGGMMKRRKLAESEENSATAHTQTGMQQSPERFGF
EPGPEFTLQTFQKYADDFSKQYFRKDTSMDSVPSVEDIEGEYWRIVEVPTEEIEVIYGADLETGTF
GSGFPKLSPETKSDAEDKYAQSGWNLNNLPRLQGSVLSFEGGDISGVLVPWVYVGMCFSSFCWHVE
DHHLYSLNYMHWGAPKLWYGVPGKDAVNLESAMRKHLPELFEEQPDLLHNLVTQFSPSLLKSEGVH
VYRCVQHEGEFVLTFPRAYHAGFNCGFNCAEAVNVAPIDWLPIGHNAVELYREQARKITISHDKLL
LGAAREAIRAQWDILFLKRNTADNMRWKSICGADSTIFKALKARIETELVQRKTLGVPAQSRKMDA
EFDSIDRECALCYYDLHLSASGCPCCPEKYACLVHAKQLCSCDWDKRFFLFRYDVNELNILADALG
GKLSAIHRWGVSDLGLSLSSCVKREKVQDSKTVRRLTDGPRRSYMSQASAVSLVSSSTSNEQKDEG
NKIMKIASPQTNNVCPSVEQRKSENISPLKEPCVRNELSCTTNSDSNGLQYNGGLGGHKGSAPGLP
VSSSPSFSSNVATRPISTSSVSMKIVQGLVASKSCIQASSRTGDSRSLLGEHHNRSPAMIHDGTNM
KSSLESSNNSCRLIASDYNATPCHSSKDQVLVTPGTNASVVTLKDSSQVHSASSQQFVRTGPWTQS
ASHEASSPSTSALKPSLDPPAMKNLYGGFTQGSAHPGPPSFSNQQPNDGRLQRTSESLPGVEARAR
GHPTVTAQPALEIHSRNGGAQKGPRIANVVHRFKCSVEPLEIGVVLSGRLWSSSQAIFPKGFRSRV
KYFSIVDPIQMAYYISEILDAGMQGPLFMVKLENCPGEVFINLSPTKCWNMVRERLNMEIRRQLNM
GKSNLPTLQPPGSVDGLEMFGLLSPPIVQAIWARDRDHICTEYWRSRPHVLIEDPNNRHMLSQGPP
LLALRGLIQRANRDELQVLRSLMTNSNNLDDSSRQQAAHIIEEEIAKQLC

SEQ ID NO: 157, DNA - *Oryza sativa*

ATGGTTTCCTCCCGCGACCCCGGCGAGGAGGCCAGCGCGCCGCCGCCCCCGCCCCCGCGCCGCGGC
GAGAAGCGGCGAATGCGCGGCCGCACCCCGTCGCCGGAGCCGGCCTCCGCGCCGCAGGATCTCTGC
CCATCAGGAGCTTGCGGGGACAATGTTGCTGGAGCTACAACTACAAATGGAAAGTGGCATCCACAT
GAATCGTACAGACCTGAAATTGATGATGCCCCTGTTTTCACTCCAACGGAAGAGGAGTTTAAAGAT
CCAATTAGATATATTACGAGCATTCGTCCCCAAGCAGAAAAGTATGGAATTTGTCGTATTGTTCCA
CCATCTTCTTGGCGACCGCCTTGTTCTCTGAAGGAGAAGAACTTCTGGGAATGTACAGAGTTCAAT
ACCCGTGTTCAACAAGTTGACAAGCTTCAAAACCGGGAACCCACAAAGAAAAAATCACAACCTCGA
GTTCAGAAGAAGAGGAAGAGGAGAAAGAGACTGAGATTTGGGATGACTCACAGGCGTCCTAGTGCA
AATACATCAGAAGACTGCGCAGATGCAGACGAGAAGTTTGGCTTTCAATCTGGCTCAGATTTCACA
CTAGATGAGTTTCAGAAATATGCAGATGAGTTTAAGCAGCAGTATTTTGGAATAAAGGGAAGTGAC
GAAATCCCTCTTTCTGAAATTAAAAAGAAGAAAAAAAATTGGCAACCATCGGTCGATGAAATAGAG
GGAGAATATTGGCGGATAGTTGTATGCCCCACTGACGAAGTTGAGGTGGATTATGGTGCTGATTTG
GACACTTCAATGTTCAGTAGTGGATTCTCTAAATTATCTTCAGATTCAAATAGACGAGATCCATAT
GGTTTATCTTGTTGGAATTTGAACAATCTTCCACGTATTCCTGGGTCTGTACTGTCATTTGAAACT
GAGGATATATCTGGCGTCGTAGTCCCTTGGCTTTATGTAGGGATGTGCTTCTCATCATTCTGTTGG

FIGURA 10 CONT.

CACGTGGAAGATCATTTCCTTTATTCTATGAATTACATGCATTTTGGTGAACCAAAAGTATGGTAT
GGTGTTCCTGGTGCTGATGCAGTGAAGCTGGAAGAAGCTATGAGAAAGAACTTACCAAGATTGTTT
GAAGAACAGCCTGATCTCCTACATGAGCTGGTTACGCAATTATCTCCTTCTGTTCTTAAATCAGAA
GGAGTTCCTGTTTATCGTGTTGTTCAGAATCCAGGCGAGTTTGTTCTAACGCTACCGCGAGCTTAC
CATTCTGGGTTCAACTGTGGCTTCAACTGTGCGGAGGCAGTAAATGTCGCACCTGTGGATTGGCTG
CCTCACGGACAATGTGCTGTTGAGCTCTACAGGGAGCAGCGGCGCAAGACATCCATATCACATGAC
AAATTATTACTAAAAACTGCAAATGAAGCTGTCAGACAGCTTTGGATGAACCTTAGCGACTGCAAA
AGTGAACAAGGAGTATACAGATGGCAGGATACTTGCGGAAAGGACGGAATGCTGACAAGTGCAATT
AAGACAAGGGTTAAAATGGAGAAGGCAGCACGGGGAGGGAATATGGCACTGCGATATAAGAAAATG
GATGGGGATTATGATTCAGCTGACCGGGAATGCTTTTCATGTTTTTATGATCTCCATTTGTCAGCT
GTCAGCTGCCAATGCTCCCCAAATCGTTTTGCTTGCTTAAACCATGCAAACATTCTATGTTCATGT
GAAATGGACAGAAAAACCGCGTTGTTGCGGTATACCATAGAGGAGCTCCATACTCTTGTTGCAGCT
CTAGAGGGTGATCCAACTGCGGTCTACCAGTGGGGACAGAATGATTTAGGTTTAGTCTGCCCATCT
GGTTCTACTCAGTACAAGAAGATGGACTTGGGTGAAAACACGGAATTTCCGGATTCAGCAACCAAC
GTCAATCATGGCTGCAGCTTAGGAAGTCAAGATCAATATCACTATGACCCCGCAAAGCCAGCAGGA
TACCAGCAAGAGAAGGGAATCCAGATTGCTTCAGAAAAACATGATAAGAACAAGATGGTTGTCAAT
CTTGAGTCTCCAGCAACAGCTAGTAATCCAAGCAGGTCAAAGTCTGACTGCAGTGGCTCACTGTCC
TTGAATCATTCATCTGAGTTACCATCTTCAAGAATTCAAACAGGAAATTCTACGCTAGCTTCCATT
ACCACAGAGAAACTGTTTGGTGTTGACATTAAATCCAATTTAGCACAGTCTTCTGATGGCCAAGTT
AGTCAATTGGCCAAGCCTTCCTCGAGCCAAACTGATGAAGTCTCTAAGCCAGCAATAGCTAAGTAT
ACGGTTGAGCTGCTAGACAGTGGAACAATGATGATTGGTAAAAAGTGGTGCAATCAGCAAGCTATA
TTCCCCAAAGGATTTAAGAGTCGAGTTACATTTCATAGTGTACTAGATCCAACAAGGACATGCTGC
TACATCTCCGAAGTTCTTGATGCTGGGCTTCTTGGACCATTGTTTAGGGTGACTGTCGAAGGTCTT
CCAGAAGTTTCGTTTACTCACACATCACCAATGCAATGTTGGGACAGTGTAAGAGACAGAGTAAAT
GAAGAAATAGCAAAACAAATAAGTTTTGGAAAATCTGGCCTTCCTGATTTTCTATCCTGCAATTCT
TTGAATGGACTTGAAATGTTTGGGTTCTTATCCTCCCCTATAATTAAGGAAATCGAGGCTCTAGAT
CCCTGTCACCAATGCTTGGACTATTGGTTGTCAAGGGTTTCTTCTGTTGGAACTGAACTCCCCTCG
GAATCTGTGATGGCAGCAATGGTTAATGACTCCACTAACCCCCCAATAAAGTTGCTCGGGATTGAG
ATTAACCGGAGGGAATCAGAACAATCAAGTAGCTTCAATAATTCCTGTGTGAGGAGGTCACACTTG
GCAGGTTGCTGA

SEQ ID NO: 158, Proteína - *Oryza sativa*

MVSSRDPGEEASAPPPPPPRRGEKRRMRGRTPSPEPASAPQDLCPSGACGDNVAGATTTNGKWHPH
ESYRPEIDDAPVFTPTEEEFKDPIRYITSIRPQAEKYGICRIVPPSSWRPPCSLKEKNFWECTEFN
TRVQQVDKLQNREPTKKKSQPRVQKKRKRRKRLRFGMTHRRPSANTSEDCADADEKFGFQSGSDFT
LDEFQKYADEFKQQYFGIKGSDEIPLSEIKKKKKNWQPSVDEIEGEYWRIVVCPTDEVEVDYGADL
DTSMFSSGFSKLSSDSNRRDPYGLSCWNLNNLPRIPGSVLSFETEDISGVVVPWLYVGMCFSSFCW
HVEDHFLYSMNYMHFGEPKVWYGVPGADAVKLEEAMRKNLPRLFEEQPDLLHELVTQLSPSVLKSE
GVPVYRVVQNPGEFVLTLPRAYHSGFNCGFNCAEAVNVAPVDWLPHGQCAVELYREQRRKTSISHD
KLLLKTANEAVRQLWMNLSDCKSEQGVYRWQDTCGKDGMLTSAIKTRVKMEKAARGGNMALRYKKM
DGDYDSADRECFSCFYDLHLSAVSCQCSPNRFACLNHANILCSCEMDRKTALLRYTIEELHTLVAA
LEGDPTAVYQWGQNDLGLVCPSGSTQYKKMDLGENTEFPDSATNVNHGCSLGSQDQYHYDPAKPAG
YQQEKGIQIASEKHDKNKMVVNLESPATASNPSRSKSDCSGSLSLNHSSELPSSRIQTGNSTLASI
TTEKLFGVDIKSNLAQSSDGQVSQLAKPSSSQTDEVSKPAIAKYTVELLDSGTMMIGKKWCNQQAI
FPKGFKSRVTFHSVLDPTRTCCYISEVLDAGLLGPLFRVTVEGLPEVSFTHTSPMQCWDSVRDRVN
EEIAKQISFGKSGLPDFLSCNSLNGLEMFGFLSSPIIKEIEALDPCHQCLDYWLSRVSSVGTELPS
ESVMAAMVNDSTNPPIKLLGIEINRRESEQSSSFNNSCVRRSHLAGC

FIGURA 10 CONT.

SEQ ID NO: 159, DNA - *Oryza sativa*

ATGTCATGTTTGAGTTGGGAGCCACCAGATATGTGGTCTAAAGTACATGGCACCGGCACTAGTCCT
GAGATGAAAAACGTGAAGGCTATTGATTGTTTATCTTGCTGTGAGGTCGAGATATGCACTCAAGAC
TTCTTCAATGGGTATTATGAAGGTCGGATGTATCAAAATCTATGGCCTGAAATGCTTAAATTGAAG
GACTGGCCTACATCAAATCATTTTGAAGAGCTTTTGCCTTCCCATGGAGTTAAATACATGAATTCT
TTACCTTTTCAACCATACACAAACTTGAAGTCTGGTTTGCTGAATGTCTCAACCTTGCTTCCTGAT
GACATCTTAAAGCTTGACATGGGCCCAAAATCATATATAGCTTATGGCTATGCACAGGAACTTGGT
AGAGGAGATTCTGTTACAAAGCTTCACTGCGATTTGTCTGATGCAGTTAACGTTTTGATGCATACT
GCTGAAGTTGACCCTTCCGAGGAACAAATAGATGCAATAAAAAGTTTGAAAAGAAGACATACAGCG
CAAAATGAAAAGGAGTGTTCTGGGAATGCAGACGGAAATTATACATCTCCTAAAATCTGTGGGGAT
GCAAATGAGTTGTCCTGTCCTATAAACAGTGAGACCAACAAGGGAGGTGCTTTATGGGATATTTTT
AGGAGGGAAGATGTTCCAAAACTGAAATTGTATCTCGACAAGCATTCTAAGGAATTTCGCCATATA
TACTGTTCTGCAGTTCAAAAGGTATGTAACCCTGTACATGATGAAACATTTTATCTAACAGAAGAA
CACAAGAGAAAACTCAAGGAGGAGCATGGAATTGAGCCTTGGACATTTGTACAAAAACTTGGGGAG
GCAGTATTCATTCCTGCTGGGTGTCCTCATCAAGTACGGAATCTTAAGTCCTGCACCAAGATTGCC
TTGGATTTTGTATCACCCGAGAATGTTAAGGAGTGTCTCAGCTTAACCGAGGACTTCCGAAGACTT
CCTAAGAACCACAGGGCCAAAGAAGACAAATTAGAGCTAGGTGTGGTTCAAAATGGACCCAAGCCC
ACATACCCTCGAGAAACTATTCTAAACTCTCAATGGGACATCTTGTTGTTTGCATATCATCCAAAT
AGTTATGACAAAGTTTAA

SEQ ID NO: 160, Proteína - *Oryza sativa*

MSCLSWEPPDMWSKVHGTGTSPEMKNVKAIDCLSCCEVEICTQDFFNGYYEGRMYQNLWPEMLKLK
DWPTSNHFEELLPSHGVKYMNSLPFQPYTNLKSGLLNVSTLLPDDILKLDMGPKSYIAYGYAQELG
RGDSVTKLHCDLSDAVNVLMHTAEVDPSEEQIDAIKSLKRRHTAQNEKECSGNADGNYTSPKICGD
ANELSCPINSETNKGGALWDIFRREDVPKLKLYLDKHSKEFRHIYCSAVQKVCNPVHDETFYLTEE
HKRKLKEEHGIEPWTFVQKLGEAVFIPAGCPHQVRNLKSCTKIALDFVSPENVKECLSLTEDFRRL
PKNHRAKEDKLELGVVQNGPKPTYPRETILNSQWDILLFAYHPNSYDKV

SEQ ID NO: 161, DNA - *Oryza sativa*

ATGCAACAGGTGGAGGGCAGGAACTGTCTTCCTGCGGAGGTCAGGATTGGCCTCGAGACGCTCAAG
AGGCGCCGGCTTGAGAGGATGCGTTTGACTGCTCAGAACAATGCCGGCGACGGTCCTCCGGTGCCC
GCAAGGAGCGGTGGGGATGCGCTAAGGACTCCCGCAAACTGCGGGGTCAGGTTGCATGCTAACAAT
GGCACAGCTCTACCTAGCAGAACCACCCAGAACAAGGACCCTTTTGCAAAGCGCAGGGTGGACAAG
TTTGATATGTCTAGCCTAGAATGGATTGACAAGATCGAAGAATGCCCTGTGTACTATCCTACCAAG
GAGGAGTTCGAGGATCCCATTGGTTATATACAGAAGATTGCACCTGTGGCTTCGAAATACGGAATT
TGCAAAATCGTATCTCCAGTAAGCGCTTCTGTTCCTGCTGGTGTCGTGTTGATGAAGGAACAGCCT
GGTTTCAAGTTCATGACCAGGGTTCAGCCGCTTCGCCTCGCCAAATGGGCTGAAGATGACACGGTC
ACTTTCTTCATGAGCGAAAGAAAGTACACTTTCCGGGATTATGAGAAAATGGCCAACAAGGTGTTC
GCCAAGAAATACTCAAGTGCTAGTTGTCTCCCAGCTAAGTACGTGGAGGAGGAATTCTGGCGCGAA
ATTGCTTTTGGTAAAATGGATTTTGTTGAATATGCCTGTGATGTTGATGGTAGTGCTTTCTCCTCT
TCTCCTCATGATCAACTTGGGAAAAGCAACTGGAACTTGAAGAATTTTTCACGGCTTTCCAATTCT
GTGCTTAGACTTCTGCAGACACCAATTCCAGGAGTAACAGATCCAATGCTTTATATCGGGATGCTC
TTCAGCATGTTTGCTTGGCATGTGGAAGATCATTATTTGTACAGCATCAATTACCATCATTGTGGG
GCATTTAAGACATGGTATGGCATACCGGGTGATGCTGCTCCTGGGTTTGAAAAGGTGGCTAGCCAG
TTTGTATACAACAAGGATATTTTGGTTGGTGAAGGAGAGGATGCAGCATTTGATGTTCTCTTGGGG
AAGACAACAATGTTCCCCCCAAATGTCTTGTTAGACCACAACGTTCCTGTTTATAAAGCTGTGCAA
AAACCTGGGGAGTTTGTCATTACTTTCCCTCGTTCCTACCACGCGGGTTTCAGCCACGGCTTCAAT
TGTGGCGAGGCTGTCAACTTTGCTATCAGTGACTGGTTTCCTCTGGGTTCTGTGGCCAGCAGACGC

FIGURA 10 CONT.

TACGCGCTTCTGAACAGAACACCCTTGCTTGCACACGAGGAGTTACTTTGCCGTTCTGCAGTGCTT
CTGTCCCACAAACTGTTAAACAGCGACCCAAAATCCCTCAATAAATCTGAGCATCCACATTCACAG
CGTTGTTTGAAGTCTTGCTTTGTGCAGTTGATGCGATTCCAGAGAAACACACGTGGCCTACTTGCT
AAAATGGGCTCTCAGATACATTATAAGCCAAAAACATACCCGAATCTCTCATGTAGCATGTGTCGG
CGTGATTGCTACATTACACATGTGTTGTGTGGATGCAACTTTGACCCAGTCTGTCTTCATCACGAA
CAAGAACTCCGGAGCTGCCCTTGTAAATCTAACCAGGTTGTCTACGTTAGGGAGGACATACAGGAG
CTAGAAGCTCTATCAAGAAAATTTGAGAAGGATATTTGCTTGGATAAGGAAATAAGTGGTTTTGAC
TCATACAAGCAGGCCGAAAAGAATGAGCCATTTTTTGAGATAACTCGGAACCTCAGGAACACTGAA
GTAAATTTGATAGAGGATGCCTTCTCAGGAGCAACTGCTGCTGATGCTGCAAAGAGTTCTCCTGCA
ACGTCAACACTGACATCTTTTGCACAACATGATGTGCCTGTTCTTGCTGAAGCAATTGTCTGTGCT
AATCAAGCCGACCAATTATACTCCACCACCGAGCAAACCATCAGCTCACCTTTAGTCAAAGGAACT
GATGCTGTGGGTGCAAATTCATCCAGCATGGCTGATGCTAATAACGGAACTGGTTCTTGTAATGCT
TCAGCTGTGGAATACAGTGGAAATTCAGATTCTGAATCTGAAATATTTCGAGTCAAGCGCAGGTCT
GGCGTATCAGTAAAGCCTGCATCTGATGCCAAGACATCAAACTTGTCTGATCAACAGGTTCTCAGG
CGGTTGAAGAAGGTGCGCCCTGAAATACAACAGCACAATAAGCGACCAGAAGACTATGGTCACTGT
TCAGTTCCCTCAGGTCGTATGAGTATGAAGAATTTGAATTCATCCTCCTCATGTGGTGAAGAACAC
TGGAGGATGAAGCGGCGGCAGTTGGAGACTCAGCAGGATGAGAGCAGTTATTCTGCAAAGCAGAAG
TCGTACTCGTATCCATCCACCAGCTATTCTTTCCGAGGAGAGTTTGTGGAAATGAGTAGAGATGCT
GCTGCAGAAGTCCGACCAAAGCGACTGAAAATCCGGCTACCTTCTTCTAGCACGAACAGAGTGGTT
GAGCAGGGCAGTTCAGGGCAAAGATTTACAAGGGATGACAAGTCGCTTGGTTGTTGGCCTGCAATT
TAG

SEQ ID NO: 162, Proteína - *Oryza sativa*

MQQVEGRNCLPAEVRIGLETLKRRRLERMRLTAQNNAGDGPPVPARSGGDALRTPANCGVRLHANN
GTALPSRTTQNKDPFAKRRVDKFDMSSLEWIDKIEECPVYYPTKEEFEDPIGYIQKIAPVASKYGI
CKIVSPVSASVPAGVVLMKEQPGFKFMTRVQPLRLAKWAEDDTVTFFMSERKYTFRDYEKMANKVF
AKKYSSASCLPAKYVEEEFWREIAFGKMDFVEYACDVDGSAFSSSPHDQLGKSNWNLKNFSRLSNS
VLRLLQTPIPGVTDPMLYIGMLFSMFAWHVEDHYLYSINYHHCGAFKTWYGIPGDAAPGFEKVASQ
FVYNKDILVGEGEDAAFDVLLGKTTMFPPNVLLDHNVPVYKAVQKPGEFVITFPRSYHAGFSHGFN
CGEAVNFAISDWFPLGSVASRRYALLNRTPLLAHEELLCRSAVLLSHKLLNSDPKSLNKSEHPHSQ
RCLKSCFVQLMRFQRNTRGLLAKMGSQIHYKPKTYPNLSCSMCRRDCYITHVLCGCNFDPVCLHHE
QELRSCPCKSNQVVYVREDIQELEALSRKFEKDICLDKEISGFDSYKQAEKNEPFFEITRNLRNTE
VNLIEDAFSGATAADAAKSSPATSTLTSFAQHDVPVLAEAIVCANQADQLYSTTEQTISSPLVKGT
DAVGANSSSMADANNGTGSCNASAVEYSGNSDSESEIFRVKRRSGVSVKPASDAKTSNLSDQQVLR
RLKKVRPEIQQHNKRPEDYGHCSVPSGRMSMKNLNSSSSCGEEHWRMKRRQLETQQDESSYSAKQK
SYSYPSTSYSFRGEFVEMSRDAAAEVRPKRLKIRLPSSSTNRVVEQGSSGQRFTRDDKSLGCWPAI

SEQ ID NO: 163, DNA - *Oryza sativa*

ATGCCGAGCGCCTTCCACTCCCTCCTCCTCCCCGCCATTCGCAACCCCAAACCTAGCCGTCGCCGC
GGCCGCGGCCGCGGCGGCAGCAAACGCCCCAAGAAGACCACCAAATCCAAGAACCGCCTCGCCGAC
GCCGCCGCCGGAGACGCCACCGCCTTCCACCTGAAGACCTCCGCGCGCGCCGGTCCGGGGGGTGCC
GGGAGCGGCCGGCGAGGCGATGGGGGATGCCTCGTGCAGCCGCTCGGCAACCTCCTCCTCCTCGGC
GGCGGCGGCAACCTCCGCGACGCGGGGCTCGGCGCGCTCCGCCCGCTCCCCGACGACGTCCTCCTC
GACGTGCTCGGCCTGCTCGCGGCGCGCGACCTCGCTAGGCTCTCCGCGGCGTCGAGGGCGCTCTAC
GTCGTCGCCTCCCACGACCCGCTCTGGCGGGCGCTCGTCCTCGACGAGCTCGGCGGGGACTTCGCC
TTCTCGGGCTCGTGGCGCGCCACCTACATCGCCGCCGCGTCGGGCGGCCGCGCCCACCTCCCCCCG
CGGGGCCTAGAGATCAGGGGGTTCTACTCCGACTACCTCTTCCAGAGCTGGCTCTGCGCCAACATG
GAGATGCGGCCGGAGTGGCTCCACCGGGACACCATCGATCGCCGCCGCGGCATGTCCGTCGAGCAA

FIGURA 10 CONT.

TTCGTCTCCGAATTCGAGGAGCCCAATAGGCCGGTGCTTCTGGAAGGCTGCCTCGAGAGCTGGCCA
GCATTGCAGAAGTGGACCAGGGAGCACTTGCTGAAGGTCTCGGCCGGGAAGGAGTTCGCCGTCGGG
CCGGTGAGCATGACGCTGGATAGGTACCTCCAGTATGCCGACAATGTGCAGGAGGAGAGGCCATTG
TACCTGTTCGATGCCAAGTTCACCGAGAAGGTGCCGGAGATGGGGAGGGACTATGAGGTGCCGGCG
TACTTCCGAGAGGACCTATTCGGGGTGCTTGGGGAGGAAAGGCCGGACCACCGATGGGTTATCATT
GGGCCGGCAGGTTCAGGGTCGTCGTTTCATGTTGATCCAAACTCGACATCGGCGTGGAATGCTGTG
ATCAAGGGAGCCAAGAAGTGGGTGATGTTCCCACCGGAGGTGGTGCCGCCAGGAGTTCATCCAAGT
GCGGATGGAGCAGAAGTCACTAGCCCTGTATCTATCATGGAATGGTTCATGAATTTTTATGGGGCA
TGTAAGACCTGGGAAAAGAGGCCTGTTGAGTGTATATGCCGGGCTGGAGAGGTGGTTTTTGTGCCT
AATGGATGGTGGCATTTGGTTATCAATCTGGAGGAATCTATTGCGATTACTCAGAATTATGTGAGC
AGGAGGAATCTGTTGAATGTTCTTGACTTTCTCAAGAGGCCCAATGCAAGTGAACTTGTATCAGGG
ACTACAGACAGGGTAAACCTGCATGACAAGTTCCGCAATGCCATCGATATGACTTATCCTGGGATG
ATTAAACAGCTTGAACTTGAAGCTCAGCAGAAGGCTGCTGCTCGCAAGAAGAAAGTCTCGTTCTGG
GAATCTGCGGTAGATGCCAATACTGGAGGATTCAAGTTCTCATTCTGA

SEQ ID NO: 164, Proteína - *Oryza sativa*

MPSAFHSLLLPAIRNPKPSRRRGRGRGGSKRPKKTTKSKNRLADAAAGDATAFHLKTSARAGPGGA
GSGRRGDGGCLVQPLGNLLLLGGGGNLRDAGLGALRPLPDDVLLDVLGLLAARDLARLSAASRALY
VVASHDPLWRALVLDELGGDFAFSGSWRATYIAAASGGRAHLPPRGLEIRGFYSDYLFQSWLCANM
EMRPEWLHRDTIDRRRGMSVEQFVSEFEEPNRPVLLEGCLESWPALQKWTREHLLKVSAGKEFAVG
PVSMTLDRYLQYADNVQEERPLYLFDAKFTEKVPEMGRDYEVPAYFREDLFGVLGEERPDHRWVII
GPAGSGSSFHVDPNSTSAWNAVIKGAKKWVMFPPEVVPPGVHPSADGAEVTSPVSIMEWFMNFYGA
CKTWEKRPVECICRAGEVVFVPNGWWHLVINLEESIAITQNYVSRRNLLNVLDFLKRPNASELVSG
TTDRVNLHDKFRNAIDMTYPGMIKQLELEAQQKAAARKKKVSFWESAVDANTGGFKFSF

SEQ ID NO: 165, DNA - *Oryza sativa*

ATGGTGGCAGACAGCGCGGTCGGCGTGCACGGCAGTGCGGCCGACAGTGGGTCGGTGCGGTCGGCG
CGGACGGCGGCAGAGAGGGACGCGGAGGCGGAGGCTACGCGGGACACAAAGGCAGCGGCGGTGGAG
CCGCCAGAGTGGCTGCAGACCCTGCCCGTGGCGCCGGAGTACCACCCGACGCTGGTGGAGTTCGCC
GACCCCATCGCCTACATCCTCAGGATCAAGCCCGAGGCGTCCCGCAACGGCATTTGCAAGATCCAG
GTCGGCCTCTCCACCAAGAACCGCCGTGCCGCCAGCCGCCGAGTGTGGGAGAGCGGGGAGCGCTAC
ACCCTGGAGGCGTTCTGCGCCAAGGTGCCCGAGTTCGAGCCCTCGAGGCACGCCGCACTGCCCAAG
AACCCCACCCACCTCCAGCTCAAGGCCCTCTTCTAG

SEQ ID NO: 166, Proteína - *Oryza sativa*

MVADSAVGVHGSAADSGSVRSARTAAERDAEAEATRDTKAAAVEPPEWLQTLPVAPEYHPTLVEFA
DPIAYILRIKPEASRNGICKIQVGLSTKNRRAASRRVWESGERYTLEAFCAKVPEFEPSRHAALPK
NPTHLQLKALF

SEQ ID NO: 167, DNA - *Oryza sativa*

ATGTCGCTGCAGCCGCCGGCGGTGGAGCCGCCGGAGTGGCTGCGGACGCTGCCCGTGGCGCCGGAG
TACCACCCGACGCTGGCGGAGTTCGCCGACCCCATCGCCTACATCCTCAGGATCGAGCCGGAGGCG
TCCCGCTACGGCATCTGCAAGATCGTGCCCCCGCTCCCGCGGCCCCCCGAGGACGACACCTTCCGC
CGCCTCCAGGCCGCCTTCGCCGCCGCGGCCTCCTCCAATGGCGACCCCTCCCCGACCTTCCCCACG
CGGCTCCAGCAGGTCGGCCTCTCCGCCAGGAACCGCCGCGCCGCCAGCCGCCGGGTGTGGGAGAGC
GGGGAGCGCTACACCCTGGAGGCGTTCCGCGCCAAGGCGGCCGAGTTCGAGCCCCCGAGGCACGCC
GCGCCGCCCAGGAACCCCACCCACCTCCAGCTCGAGGCCCTCTTCTGGGCCGCCTGCGCCTCCAGG
CCCTTCAGCGTCGAGTACGGCAACGACATGCCCGGCTCCGGCTTCGCCTCCCCCGACGAGCTCCCC

FIGURA 10 CONT.

GACGCCGCCAATGCCACCGACGTCGGGGAGACGGAGTGGAACATGCGGGTGGCGCCCCGCGCCCGG
GGGTCGCTGCTCCGCGCCATGGCCCGCGACGTGGCGGGCGTCACCACCCCGATGCTGTACGTGGCC
ATGCTCTACAGCTGGTTCGCCTGGCACGTGGAGGACCACGAGCTCCACAGCCTCAACTTCCTCCAC
TTCGGCAAGGCCAAGACCTGGTACGGCGTCCCCCGCGACGCCATGCTCGCCTTCGAGGAGACGGTG
CGCGTCCACGGCTACGCCGACGACCTCAACGCCATCATGGCCTTTCAAACGTTAAATGAGAAGACA
ACTGTCTTATCTCCTGAGGTGCTTCTTTCTGCTGGTGTTCCCTGCTGCAGATTGGTTCAAAAAGCG
GGAGAATTTGTTATCACATTCCCTGGAGCTTATCATTCAGGCTTTAGCCATGGATTTAATTGTGGG
GAAGCATCAAATATTGCAACTCCCCATTGGTTACAAGTGGCTAAAGAAGCTGCAATCCGGAGGGCT
TCAACTAATTGTGGTCCGATGGTGTCTCATTATCAGCTGCTTTATGAGCTAGCACTCTCATTGCGT
CCAAGGGAGCCTAAGAATTTCTATTCTGTTCCAAGAAGCTCACGTTTAAGGGACAAGAATAAGAAT
GAAGGTGATATAATGGTCAAAGAAAATTTTGTTGGAAGTGTAACCGAAAACAACAATTTGCTTAGC
GCCCTTCTGGATAAGAATTCTTGTATAATTGTTCCAAACGCTGATTTCTTCGTTCCATCCTTTCCT
GTTGCACTGGAATCTGAAGTTACTGTTAAACAAAGGTTTACAGCTGGTCCCTGCAGTATTAGCCAG
CAAGGAGCAGAAAATATGGCTGCTGATCATGTGGCTGTAGACAAGGTTACAGAGATTCAGGACATG
AGTGGATCCTTATATCCTTGTGAAACCAGTCTTGTGGGCTGTAGCAATAGAAAACTTTATGAAACA
AAATATGGCCAACGGGATGCTGCTGCTTTGTGCTTATCTACTTCAGAGATTCAGAGCAGAGGAATT
GATACAGCAAGATCACATCCAGCAGGCGGGATCTTAGATCAAGGACGGTTGCCATGTGTACAGTGC
GGAATACTAAGCTTTGCATGTGTAGCCATCATTCAACCTAGAGAAGCAGCAGTACAGTTTATCATG
TCTAAAGAGTGCATCTCATCGAGTGCAAAACAAGGAGGAATTGGTGCATCTGATGATACCTCAAAC
TGGATTGACCAGAGTCATGAGATTAGTCCTCCACCAGGTCCAGCATCTGGAACAGATGATAATGTA
AAGCATGCCGTAAGTTTGGCCCATGTCTCTGATCGGTGTAGAGAACTATATGCCAGCAATACAGAT
GGATGCACCTCTGCTCTTGGGCTTCTAGCTTCTGCATATGACTCATCAGACTCTGATGATGAAACA
ACTGAGGATGTTTCAAAACATAGCAAGAAAAATGATTCAGTAAACCAAAGCACGGATCCTCAAATC
TTGGAAACATCAGCTAGCTGTTCCAGTACAGTTCAGTGTCAAAAGACAAATTCACACTTGCATGAA
GAGGAATGTGAGGCAAGAGCCACATCATTGATGAAACCTGTAAGCCATAATAGTAGGCCCATTAGT
CAGTCTAACAGGGATACAGATATCGACCATTTTATAGAACTGGGAAAGTCAGGAACACAGTGTTCA
GGTTATCTTGATTTGGTTGATGATCTAACTACATCAGTTTTAAAATCCTCTTCAGATACTTGTGTA
AGTGCAGCTAAAGCCTCAATGGATCCAGATGTTTTGACCATGCTCAGGTACAATAAAGACTCTTGC
AGAATGCATGTGTTTTGTCTTGAACATGCTTTGGAAACATGGACACAGCTTCAACAAATCGGTGGT
GCTAATATTATGCTTTTGTGCCATCCAGAATACCCTAGAGCAGAGTCAGCAGCAAAAGTCATAGCA
GAAGAACTTGGTATCAAGCATGATTGGAAAGATATCACTTTCAAAGAAGCAACTGAAGAGGACGTT
AAAAAGATTCAGTTGGCTTTACAGGATGAAGATGCTGAACCCACTGGCAGTGACTGGGCAGTCAAA
ATGGGTATCAACATATATTACAGTGCCAAACAGAGCAAATCACCACTTTACAGCAAGCAGATACCA
TACAATTCCATCATTTATAAGGCATTTGGCCAAGAAAACCCAGACAGTTTGACAGATTATGGATGT
CAAAAGTCAGGCTCAACAAAGAAAAAGGTGGCTGGGTGGTGGTGTGGGAAAGTTTGGATGTCAAAT
CAAGTTCATCCACTTCTAGCTCGTGAGCGTGAAGAACAGAACAGTAGCGTAGTGTACGGCAAAGCA
ATGTTCACTACTATTTCCCATGGCAAAGTACAAGATGAAGCGTCAACAAGATGCAATACAAGCAAC
CGAACCCCATCAAGAAGAACATCAAGAAGAAAAAAGGGGGTATCTGCTGAGAAATCTAAACCAAAA
AACAAGAGATCCACTGCTTCTGATGAAGCTAGTATGCTTTGCAGTGGCCTTGGAATGAACTCTGGA
GTTATCCATGACCAAACTGAAAACTCTGATGATTATGACAAACATGGCAATGGAGATGAAATTGAG
GAAGGAACAAATCCTCAGAAATATCAACAACGTAAATTGCAAAACGTGACCAGGAAATCAAGTTCT
AAGAAGCGGAAGGATGAGAAAAGAACAGATAGTTTTCATGAACTATATGACGAGGATAATGGTGTG
GATTATTGGCTTAACATGGGTAGTGGAGATGATGCCACACTAGGTAACTCTCGACAACAAAGTCCT
GATCCAGTGAAAGTGAAATCTGGAGGCAAATTGCAAGGTAAGAGAAAATCTAGCAAGTACAAGTCC
AATGATGATTTGTTGAATGAAGAAAATAAGTTACAAAAGATGAACAAAAAATCAAGCTCTAAGAAG
CAAAAGAATGATAAGATAAATAGACAACTTCAAGAAGATCAAACCGAAGATGACCATATGGACCAC
TTAGTTGACGTAGCGGTTGCAGATGAAGTCACACTAGACAATGAGGATAAAATTACAGAAGACAAA
ATTGATGACGTGAAAGTGAAATCTAGAGGAAAATCGCAAAATGGTAAGAGAAAAGGCAGCAAACAT

FIGURA 10 CONT.

CAGGCTACCGATGGTTTGCGCGCTGGAAATAAAGTGGCCAAATTTCCGTGTGATATAGAAGGATGT
GACATGAGCTTCAGTACCCAGCAGGACTTGTTATTGCATAAGCGTGACATTTGTCCTGTCAAAGGA
TGCAAGAAGAAGTTCTTCTGCCACAAGTACTTGCTTCAGCACCGGAAGGTGCATATAGATGAAAGG
CCACTTAAGTGTACATGGAAGGGGTGTAAGAAGGCATTCAAGTGGCCATGGGCGAGGACGGAGCAC
ATGAGAGTCCATACGGGGGTAAGACCTTACGAGTGCCAGGAACCTGGCTGTGGTCAGACATTCCGA
TTCGTTTCAGATTTCAGCCGCCACAAGAGAAAGACTGGCCATTCCTCTGATAAGAGAAGAAAGAAC
AGTACATAA

SEQ ID NO: 168, Proteína - *Oryza sativa*

MSLQPPAVEPPEWLRTLPVAPEYHPTLAEFADPIAYILRIEPEASRYGICKIVPPLPRPPEDDTFR
RLQAAFAAAASSNGDPSPTFPTRLQQVGLSARNRRAASRRVWESGERYTLEAFRAKAAEFEPPRHA
APPRNPTHLQLEALFWAACASRPFSVEYGNDMPGSGFASPDELPDAANATDVGETEWNMRVAPRAR
GSLLRAMARDVAGVTTPMLYVAMLYSWFAWHVEDHELHSLNFLHFGKAKTWYGVPRDAMLAFEETV
RVHGYADDLNAIMAFQTLNEKTTVLSPEVLLSAGVPCCRLVQKAGEFVITFPGAYHSGFSHGFNCG
EASNIATPHWLQVAKEAAIRRASTNCGPMVSHYQLLYELALSLRPREPKNFYSVPRSSRLRDKNKN
EGDIMVKENFVGSVTENNNLLSALLDKNSCIIVPNADFFVPSFPVALESEVTVKQRFTAGPCSISQ
QGAENMAADHVAVDKVTEIQDMSGSLYPCETSLVGCSNRKLYETKYGQRDAAALCLSTSEIQSRGI
DTARSHPAGGILDQGRLPCVQCGILSFACVAIIQPREAAVQFIMSKECISSSAKQGGIGASDDTSN
WIDQSHEISPPPGPASGTDDNVKHAVSLAHVSDRCRELYASNTDGCTSALGLLASAYDSSDSDDET
TEDVSKHSKKNDSVNQSTDPQILETSASCSSTVQCQKTNSHLHEEECEARATSLMKPVSHNSRPIS
QSNRDTDIDHFIELGKSGTQCSGYLDLVDDLTTSVLKSSSDTCVSAAKASMDPDVLTMLRYNKDSC
RMHVFCLEHALETWTQLQQIGGANIMLLCHPEYPRAESAAKVIAEELGIKHDWKDITFKEATEEDV
KKIQLALQDEDAEPTGSDWAVKMGINIYYSAKQSKSPLYSKQIPYNSIIYKAFGQENPDSLTDYGC
QKSGSTKKKVAGWWCGKVWMSNQVHPLLAREREEQNSSVVYGKAMFTTISHGKVQDEASTRCNTSN
RTPSRRTSRRKKGVSAEKSKPKNKRSTASDEASMLCSGLGMNSGVIHDQTENSDDYDKHGNGDEIE
EGTNPQKYQQRKLQNVTRKSSSKKRKDEKRTDSFHELYDEDNGVDYWLNMGSGDDATLGNSRQQSP
DPVKVKSGGKLQGKRKSSKYKSNDDLLNEENKLQKMNKKSSSKKQKNDKINRQLQEDQTEDDHMDH
LVDVAVADEVTLDNEDKITEDKIDDVKVKSRGKSQNGKRKGSKHQATDGLRAGNKVAKFPCDIEGC
DMSFSTQQDLLLHKRDICPVKGCKKKFFCHKYLLQHRKVHIDERPLKCTWKGCKKAFKWPWARTEH
MRVHTGVRPYECQEPGCGQTFRFVSDFSRHKRKTGHSSDKRRKNST

SEQ ID NO: 169, DNA - Glycine max

ATGTTGACGACAACTGTATCCGGCGGCGACCCTCCTTCTCGCGGTTTTGATACGCCGACGCTGGAC
CGGGAAGCGGCGGCGCTGCTCCACGCGATCTCCGAGCATGGCGGGTACGCGTACGTGAGCATGGCG
GTGCTGGCTTCCGGCGGCGACATTCGCGCGGCGGAGGCGGCGTGGGAGATGGCGTGGGAGCAGCTG
CACTCGGGTCCGTGGCACTCGGTGCTGCCGGTGTGGCGCGACGCTTACTCCATGGCGTGCCTCCAC
GTGGCGCGCCACCACTACGGCAACGGTGAGTTCTTGGACGCACTTAGGGTTTTGGATTTGGGAATC
ATCATGGGAGGCACGCTCCTCCGCAAGGATTTGGACTCCGCCATCGAGAAAATGTCGGAACAAACG
AGGAAGACCGTTAGGGTTTCTGATTTGGGGAACTCCGAGCACCGACTCGTCGATCGCGAATTTGAT
ATGGCAGAGGTGCTCCAACTTTTACCTGTGAAGTCTCTTTCAACGAAACTTGTGGTGAAGAAATCG
GCGCTGTCCTTGGAGAAATTCCTGAAGGATCATTACCTGTCTGGCTGCCCGGTTATTATCAGTGAT
TGTATGTCTCACTGGCCAGCCAAGATGAAATGGAATGACGAAGATTACTTGCTGAGAGTTGCCGGA
GACCGTACAGTTCCAGTTGAGGTTGGGAAAAACTATTTATGTACTGAGTGGAAGCAAGAGCTAATT
ACTTTTTCAGAGTTTCTTCAGCGGATAAAGTCTGATAGCTGTTCTCCTGGTGGTCCTACATATCTT
GCTCAGCATCCATTATTTGATCAGATAAATGAGCTTCGGAAAGATATCTTTATTCCTGACTATTGT
TTTACTGGTGGAGGGGAGCTACGATCTCTCAATGCTTGGTTTGGTCCAGCAGGAACAGTAACACCG
TTACACCATGATCCACATCATAACATACTAGCTCAGGTTGTTGGAAAGAAATACATTAGGCTATAC

FIGURA 10 CONT.

TCTTCGTCTTTATCTGAGGAACTTTCCCCCCACTCTGGTACCATGCTCCACAACTCCAGCCAGGTT
GATTTAGATGATATGGATGAAAAGAAGTTTCCGAAGGTGCAAGACTTGGAATTTGTAGACTGTATT
TTAGAGGAAGGCGAAATGTTATATATCCCGCCAAAATGGTGGCACTATGTGCGGTCTTTGACTACC
AGTTTTTCGGTTAGCTTTTGGTGGAGTGAGGGTGAAAGTTCGGATGCATCCTAA

SEQ ID NO: 170, Proteína - Glycine max
MLTTTVSGGDPPSRGFDTPTLDREAAALLHAISEHGGYAYVSMAVLASGGDIRAAEAAWEMAWEQL
HSGPWHSVLPVWRDAYSMACLHVARHHYGNGEFLDALRVLDLGIIMGGTLLRKDLDSAIEKMSEQT
RKTVRVSDLGNSEHRLVDREFDMAEVLQLLPVKSLSTKLVVKKSALSLEKFLKDHYLSGCPVIISD
CMSHWPAKMKWNDEDYLLRVAGDRTVPVEVGKNYLCTEWKQELITFSEFLQRIKSDSCSPGGPTYL
AQHPLFDQINELRKDIFIPDYCFTGGGELRSLNAWFGPAGTVTPLHHDPHHNILAQVVGKKYIRLY
SSSLSEELSPHSGTMLHNSSQVDLDDMDEKKFPKVQDLEFVDCILEEGEMLYIPPKWWHYVRSLTT
SFSVSFWWSEGESSDAS

FIGURA 10 CONT.

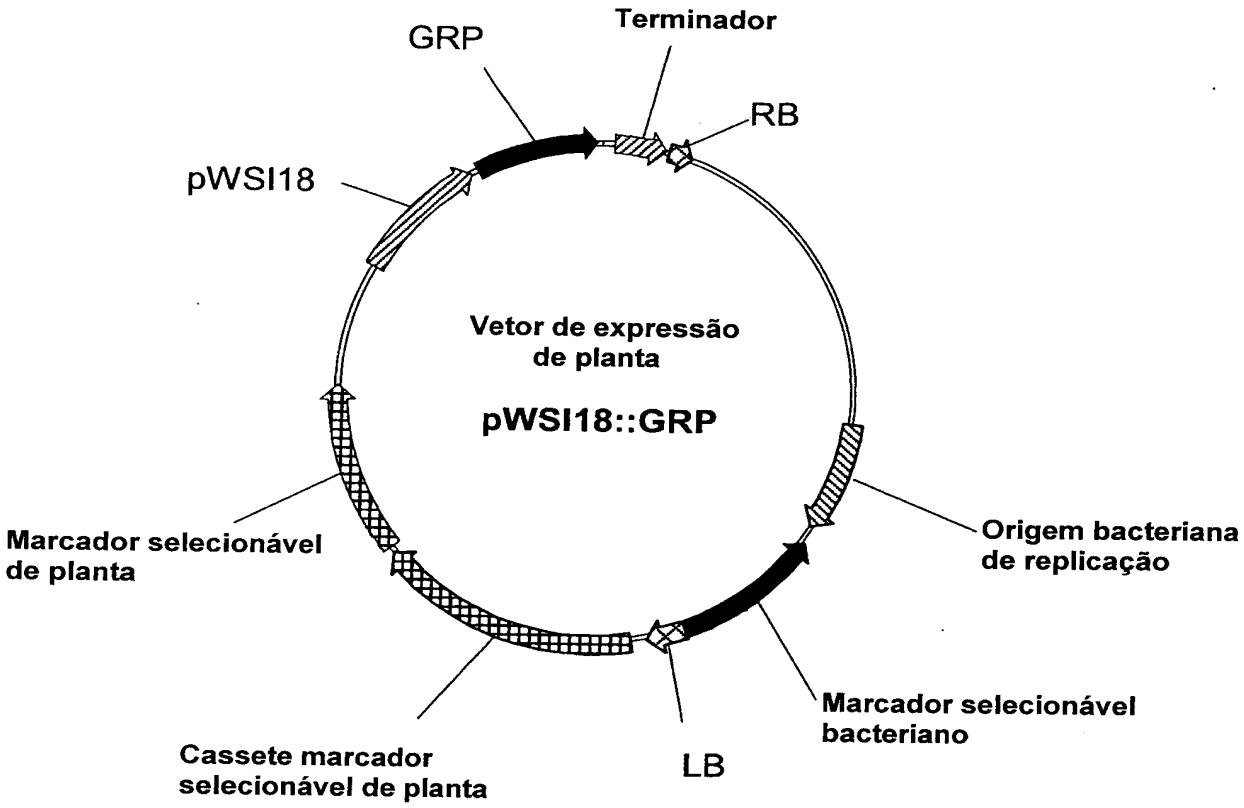

GRP
Terminador
RB
pWSI18
Vetor de expressão
de planta
pWSI18::GRP
Marcador selecionável
de planta
Origem bacteriana
de replicação
Marcador selecionável
bacteriano
Cassete marcador
selecionável de planta
LB

FIGURA 11

SEQ ID NO: 171, DNA - *Arabidopsis thaliana*

CTTAGTCAATAGCATCTTCTCCGATCATCGACGTCCGGAGATTTCTTCGGCCCTCTCATACTTGAG
TTGTTGTAGGATTTGTTGCTCTGGCTCTGGTGGTAGGTCTATGAAATCAACCCATATCGTGAATGG
ACTGCAACATGGTATCTTCGTCCCAGTGGGATTGGGAGCATTTGATCATGTCCAATCCGTCAAGGA
CTGAAGATGACAGCAAACAGCTACCTACTGAGTGGGAAATTGAAAAAGGTGAAGGAATTGAATCTA
TAGTTCCACATTTCTCAGGCCTTGAGAGAGTCAGTAGTGGCTCTGCCACCAGCTTCTGGCACACTG
CTGTATCGAAAAGCTCACAGTCGACCTCTATCAACTCATCATCTCCCGAAGCCAAACGATGCAAGC
TTGCATCAGAAAGTTCCCCTGGAGATTCTTGCAGCAACATAGACTTTGTCCAGGTGAAGGCTCCCA
CAGCTCTCGAGGTATCCGTTGCCTCAGCTGAATCAGATCTTTGTTTAAAACTAGGAAAGCGGACAT
ACTCTGAAGAATACTGGGGTAGAAACAATAATGAAATTTCAGCGGTTTCTATGAAGTTGTTAACTC
CATCTGTTGTCGCTGGGAAATCCAAATTGTGTGGTCAGAGCATGCCAGTCCCGCGTTGCCAAATTG
ATGGCTGTGAACTGGATCTCTCATCTGCTAAGGGTTATCATCGTAAGCACAAAGTCTGCGAAAAGC
ATTCAAAGTGCCCAAAAGTTAGCGTGAGTGGCCTGGAACGTCGGTTCTGCCAACAGTGTAGCAGGT
TCCATGCTGTCTCTGAATTTGATGAGAAGAAACGAAGCTGCCGAAAACGTCTTTCTCATCATAATG
CGAGGCGTCGTAAGCCACAAGGAGTATTTTCAATGAATCCCGAGAGGGTGTATGATCGAAGACAGC
ATACAAATATGTTGTGGAATGGGGTGTCCCTTAACGCGAGATCTGAAGAAATGTATGAATGGGGTA
ATAACACTTATGATACAAAGCCTAGACAAACGGAAAAAAGCTTTACTCTGAGCTTCCAGAGAGGTA
ATGGCTCTGAGGACCAGCTGGTTGCTAGTAGCAGCCGTATGTTCTCTACATCTCAAACCTCAGGTG
GGTTCCCAGCAGGAAAGTCCAAGTTTCAACTTCATGGCGAAGATGTGGGAGAATACTCAGGAGTCC
TCCATGAATCTCAAGATATCCACCGTGCTCTCTCTCTTCTGTCAACCTCTTCGGATCCCCTGGCCC
AACCACATGTGCAGCCATTTTCTCTACTCTGTTCATATGATGTTGTACCAAAATAGATGAGTAAGT
AATGTGTAATTTGTAAACCTGTTACTCAGTTGGTGGATACTTTTCCAAACCTATGATAAAAACCTC
GTCCTAGATCCCGTTAAATGCCAAACTTTCGGCTACTATAACTATGTTATCGTTATCATTATCATT
GTTTAACACCCT

SEQ ID NO: 172, Proteína - *Arabidopsis thaliana*

MDCNMVSSSQWDWEHLIMSNPSRTEDDSKQLPTEWEIEKGEGIESIVPHFSGLERVSSGSATSFWH
TAVSKSSQSTSINSSSPEAKRCKLASESSPGDSCSNIDFVQVKAPTALEVSVASAESDLCLKLGKR
TYSEEYWGRNNNEISAVSMKLLTPSVVAGKSKLCGQSMPVPRCQIDGCELDLSSAKGYHRKHKVCE
KHSKCPKVSVSGLERRFCQQCSRFHAVSEFDEKKRSCRKRLSHHNARRRKPQGVFSMNPERVYDRR
QHTNMLWNGVSLNARSEEMYEWGNNTYDTKPRQTEKSFTLSFQRGNGSEDQLVASSSRMFSTSQTS
GGFPAGKSKFQLHGEDVGEYSGVLHESQDIHRALSLLSTSSDPLAQPHVQPFSLLCSYDVVPK

SEQ ID NO: 173, DNA - *Nicotiana tabacum*

GTACAAAAAAGCAGGCTGGTACCGGTCCGGAATTCCCGGGATATCGTCGACCCACGCGTCCGCCGT
GATGTAATCTTGGTGATGCTACTTATTCCCTTTTCCCTTCTTGAGCCCAAACTCAAGAAGGTCAAA
AACAAAAAAATTACAAAAAGCTGGAATCTTGCAGTTTTTTTATTTAATTTATTTATCCTATGTTGA
ATTAATTTTTGGGGTCAATATTTCCCAATTTGTAGTCTCCAATGGAGCCTCGTGTTGGTAATAAGT
TCCGGCTTGGCCGGAAAATCGGTAGCGGTTCTTTTGGAGAGATCTATCTCGGCGCTAATGTTCAAA
CTAACGAAGAGGTTGCAATTAAGCTGGAAAATGTGAAAACAAAGCATCCTCAACTATTATACGAAG
CAAAGTTGTATAAAATACTACAAGGAGGAACTGGAATCCCCAATTTAAAATGGTTTGGAGTTGAAG
GAGATTATAATGCCCTTGTGATGGATTTGCTGGGGCCTAGTCTTGAAGATCTCTTCAACTTCTGCA
GTAGGAAGCTGTCTTTAAAGACCGTTCTCATGCTCGCAGATCAGATGATTAATCGGGTTGAATTTG
TTCATGCCAAATCTTTTCTTCATCGAGATATAAAACCTGACAACTTTCTTATGGGATTAGGAAGAC
GTGCAAATCAGGTCTATATGATTGATTTTGGGCTGGCCAAGAAGTATAGAGACTCATCAACTCATC
AGCATATTCCGTATAGAGAAAACAAAAATTTGACAGGAACTGCTAGATACGCAAGCATGAATACTC
ATCTTGGCATTGAACAAAGTCGAAGGGATGATTTGGAATCGCTGGGTTATGTTTTAATGTACTTCT
TAAGAGGAAGTCTCCCTTGGCAGGGGCTGAAAGCAGGCACTAAGAAACAGAAGTATGAGAAGATCA

FIGURA 12

GTGAGAAGAAAGTATCAACATCAATAGAGACCTTGTGTAGGGGATATCCTGCAGAGTTTGCATCAT
ATTTTCATTACTGTCGATCACTAAGATTTGATGATAAACCAGATTATGCTTATCTGAAGAGAATTT
TCCGTGATCTTTTCATTCGTGAAGGGTTTCAATTTGATTATATATTTGACTGGACCATTTTGAAAT
ATCAGCAATCACAGCTTGCCAATCCTCCATCTCGTGCTCTTGGTGGTACTGCTGGGCCAAGCTCAG
GGATGCCTCATGCTCTTGTTAATGTTGAGAGGCAATCAGGTGGAGATGAAGGTCGACCAACTGGTT
GGTCTTCATCAAATCTTACACGTAATAAGAGCACGGGGCTGCATTTCAATTCTGGAAGCTTATTGA
AGCAAAAAGGCACAGTTGCTAATGATTTATCCATGGGTAAAGAGTTATCCAGTTCTAATTTTTTCC
GGTCAAGTGGACCATTGAGGCGTCCAGTTGTCTCTAGCATCCGAGACCCAGTGATTGCAGGGGGTG
AACCTGACCCCTCCGGCACTCTGACAAAAGATGCAAGCCCGGGACCATTGCGTAAAGTATCCAGTG
CTGCACGGAGGAGTTCACCAGTTGTGTCCTCAGATCACAAGCGCAGCTCCTCTATCAAAAATGCCA
ACATAAAGAATTTAGAGTCCACCGTCAAGGGAATAGAGGGTTTAAGTTTTCGATGATGAGGGACTG
CATTAGTAGCTGTGCTTTGTCTCAGTTCTCCGTTCACTGTAAATTTTGGCACACCAACTTGGGGAG
TAAGAGTTCTGATATTAGTTGCTGTCAGGAAGTACCATAAAGCTGAATTATACAATTAAAATTTGG
GATCCAATCGCAAAAGCACATTAAGGATATGATGGGGTTGCAGATCCAAACTCACAGATTCCAGTT
TATGCTCGTCCATACAGTTATAGGCACTTTCCATATTCTTTTCTTTAATCTCTGTCTCTTGCTTGT
TATTGTTATGTCGTGGTATTCTTGTTGAGGTCATGTTTGTGAATTGCGAAGATGGTCATGTATAAT
TGCCGAGAAATCATGTACTAGTTTGTTTTAAACATGAGCAAACTGTTATTTTGTTCAAGCTACTTT
AATATCAAAAAAAAAAAAAAAAGGGCGGCCGCTCTAGAGTATCCCTCGAGGGGCCCAAGCTTACGC
GTACCCAGCTTT

SEQ ID NO: 174, Proteína - *Nicotiana tabacum*

MEPRVGNKFRLGRKIGSGSFGEIYLGANVQTNEEVAIKLENVKTKHPQLLYEAKLYKILQGGTGIP
NLKWFGVEGDYNALVMDLLGPSLEDLFNFCSRKLSLKTVLMLADQMINRVEFVHAKSFLHRDIKPD
NFLMGLGRRANQVYMIDFGLAKKYRDSSTHQHIPYRENKNLTGTARYASMNTHLGIEQSRRDDLES
LGYVLMYFLRGSLPWQGLKAGTKKQKYEKISEKKVSTSIETLCRGYPAEFASYFHYCRSLRFDDKP
DYAYLKRIFRDLFIREGFQFDYIFDWTILKYQQSQLANPPSRALGGTAGPSSGMPHALVNVERQSG
GDEGRPTGWSSSNLTRNKSTGLHFNSGSLLKQKGTVANDLSMGKELSSSNFFRSSGPLRRPVVSSI
RDPVIAGGEPDPSGTLTKDASPGPLRKVSSAARRSSPVVSSDHKRSSSIKNANIKNLESTVKGIEG
LSFR

SEQ ID NO: 175, *Oryza sativa* – sequência de promotor de WSI18

gcttgagtcatagggagaaaacaaatcgatcatatttgactcttttccctccatctctcttaccgg
caaaaaaagtagtactggtttatatgtaaagtaagattctttaattatgtgagatccggcttaatg
cttttcttttgtcacatatactgcattgcaacaattgccatatattcacttctgccatcccattat
atagcaactcaagaatggattgatatatcccctattactaatctagacatgttaaggctgagttgg
gcagtccatcttcccaacccaccaccttcgtttttcgcgcacatactttcaaactactaaatggt
gtgtttttaaaaatattttcaatacaaaagttgctttaaaaaattatattgatccatttttttaa
aaaaaatagctaatacttaattaatcacgtgttaaaagaccgctccgttttgcgtgcaggagggat
aggttcacatcctgcattaccgaacacagcctaaatcttgttgtctagattcgtagtactggatat
attaaatcatgttctaagttactatatactgagatgaatagaataagtaaaattagacccacctta
agtcttgatgaagttactactagctgcgtttgggaggacttcccaaaaaaaaaagtattagccatt
agcacgtgattaattaagtactagtttaaaaaacttaaaaaataaattaatatgattctcttaagt
AACTCTCCTATAGAAAACTTTTACAAAATTACACCGTTTAATAGTTTGGAAAATATGTCAGTAAAA
AATAAGAGAGTAGAAGTTATGAAAGTTAGAAAAAGAATTGTTTTAGTAGTATACAGTTATAAACTA
TTCCCTCTGTTCTAAAACATAAGGGATTATGGATGGATTCGACATGTACCAGTACCATGAATCGAA
TCCAGACAAGTTTTTTATGCATATTTATTCTACTATAATATATCACATCTGCTCTAAATATCTTAT
ATTTCGAGGTGGAGACTGTCGCTATGTTTTTCTGCCCGTTGCTAAGCACACGCCACCCCCGATGCG
GGGACGCCTCTGGCCTTCTTGCCACGATAATTGAATGGAACTTCCACATTCAGATTCGATAGGTGA

FIGURA 12 CONT.

CCGTCGACTCCAAGTGCTTTGCACAAAACAACTCCGGCCTCCCGGCCACCAGTCACACGACTCACG
GCACTACCACCCCTGACTCCCTGAGGCGGACCTGCCACTGTTCTGCATGCGAAGCTATCTAAAATT
CTGAAGCAAAGAAAGCACAGCACATGCTCCGGGACACGCGCCACCCGGCGGAAAAGGGCTCGGTGT
GGCGATCTCACAGCCGCATATCGCATTTCACAAGCCGCCCATCTCCACCGGCTTCACGAGGCTCAT
CGCGGCACGACCGCGCACGGAACGCACGCGGCCGACCCGCGCGCCTCGATGCGCGAGCCCATCCGC
CGCGTCCTCCCTTTGCCTTTGCCGCTATCCTCTCGGTCGTATCCCGTTTCTCTGTCTTTTGCTCCC
CGGCGCGCGCCAGTTCGGAGTACCAGCGAAACCCGGACACCTGGTACACCTCCGCCGGCCACAACG
CGTGTCCCCCCTACGTGGCCGCGCAGCACATGCCCATGCGCGACACGTGCACCTCCTCATCCAAAC
TCTCAAGTCTCAACGGTCCTATAAATGCACGGATAGCCTCAAGCTGCTCGTCACAAGGCAAGAGGC
AAGAGGCAAGAGCATCCGTATTAACCAGCCTTTTGAGACTTGAGAGTGTGTGTGACTCGATCCAGC
GTAGTTTCAGTTCGTGTGTTGGTGAGTGATTCCAGCCAAGTTTGCG

SEQ ID NO: 176, *Oryza sativa* – sequência promotora de GOS2

AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAATAT
AAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCACCTACT
TTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGTAATTAAGT
GGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGTTAAATTATTCG
AGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTTCTAGCTGAACTCAATGGGT
AAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTCTGAACGTATTGGCAAAGATTTAAA
CATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACATCATTAAGGACATGTCTTA
CTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATTTTTATTATTTATCTTTTTTCG
ATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAAT
ACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTAGGTAGCAATATC
TGAATTCAAGCACTCCACCATCACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTT
ACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAAAAAAGAATT
TTGCTCGTGCGCGAGCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCCACA
GAACAACCCACAAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAG
GCTTTGCGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGCATCCTCCTTCTCCCATCTATAA
ATTCCTCCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAAG
GACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTCGATCCATATCTTCCGGTCGAGTTCTTGGT
CGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCTCCCTTCGGTTGTTCTTGGATTT
ATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCC
TGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTCGGTTTGATTAGTAGT
ATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTAGGGATCGGAATCTTGCGATT
TTGTGAGTACCTTTTGTTTGAGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAGTAC
GGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTC
CCTATTGAACAAAAATAATCCAACTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTT
AAGCCTGTCCAAAATTTCGCAGCTGGCTTGTTTAGATACAGTAGTCCCCATCACGAAATTCATGGA
AACAGTTATAATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTT
TTTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATG
CTTTTATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGA
AGAACTTATCCGATTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATT
CATTTGGATTATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAA
CTGTCCTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGT
AGAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAATCG
GGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCTTGCCACT
TTCACCAGCAAAGTTC

FIGURA 12 CONT.

SEQ ID NO: 177, Sequência artificial - prm86687
GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGGACTGCAACATGGTATCT

SEQ ID NO: 178, Sequência artificial - prm8668
GGGGACCACTTTGTACAAGAAAGCTGGGTCACATTACTTACTCATCTATTTTGG

FIGURA 12 CONT.

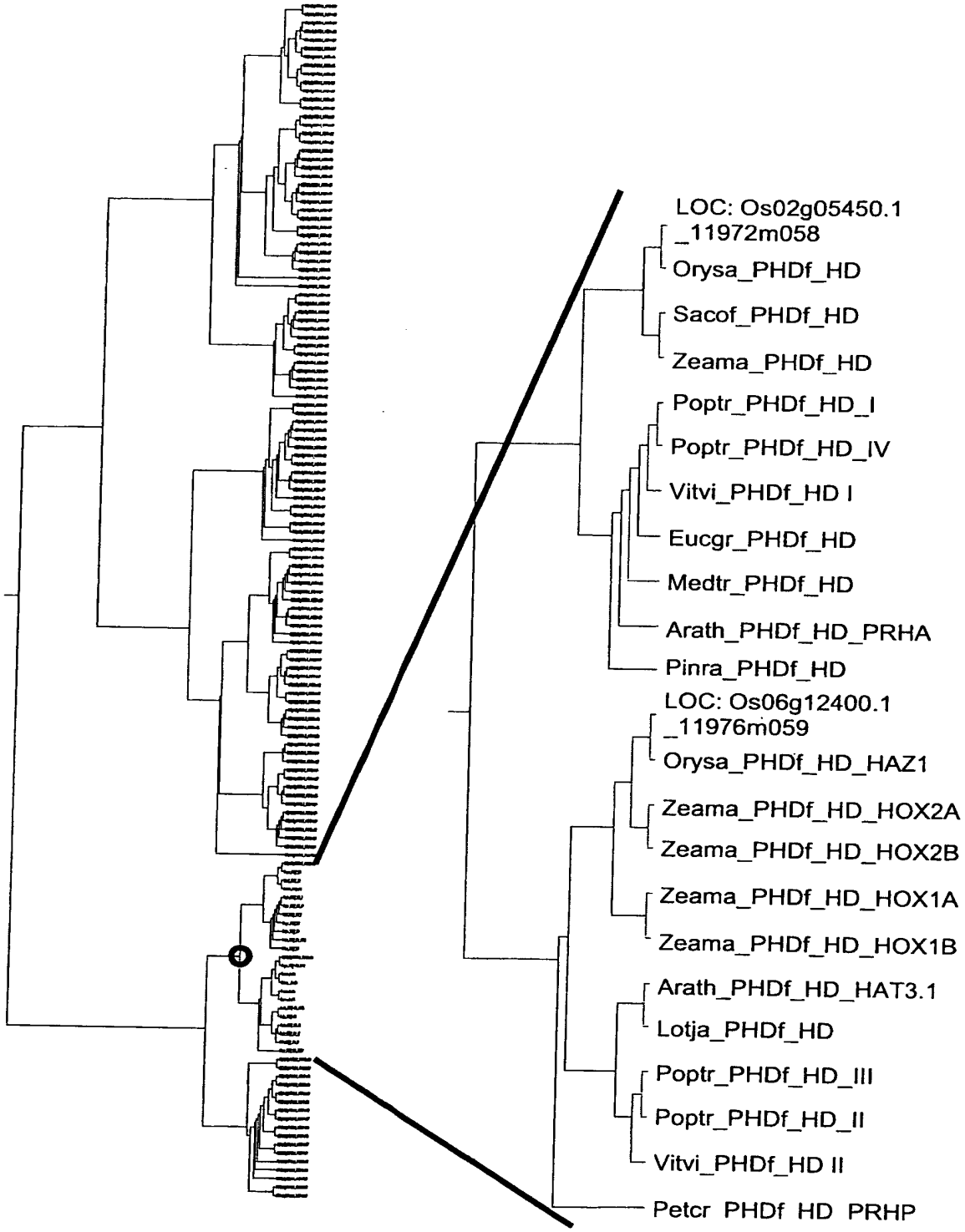

LOC: Os02g05450.1
_11972m058
Orysa_PHDf_HD
Sacof_PHDf_HD
Zeama_PHDf_HD
Poptr_PHDf_HD_I
Poptr_PHDf_HD_IV
Vitvi_PHDf_HD I
Eucgr_PHDf_HD
Medtr_PHDf_HD
Arath_PHDf_HD_PRHA
Pinra_PHDf_HD
LOC: Os06g12400.1
_11976m059
Orysa_PHDf_HD_HAZ1
Zeama_PHDf_HD_HOX2A
Zeama_PHDf_HD_HOX2B
Zeama_PHDf_HD_HOX1A
Zeama_PHDf_HD_HOX1B
Arath_PHDf_HD_HAT3.1
Lotja_PHDf_HD
Poptr_PHDf_HD_III
Poptr_PHDf_HD_II
Vitvi_PHDf_HD II
Petcr PHDf HD PRHP

FIGURA 13

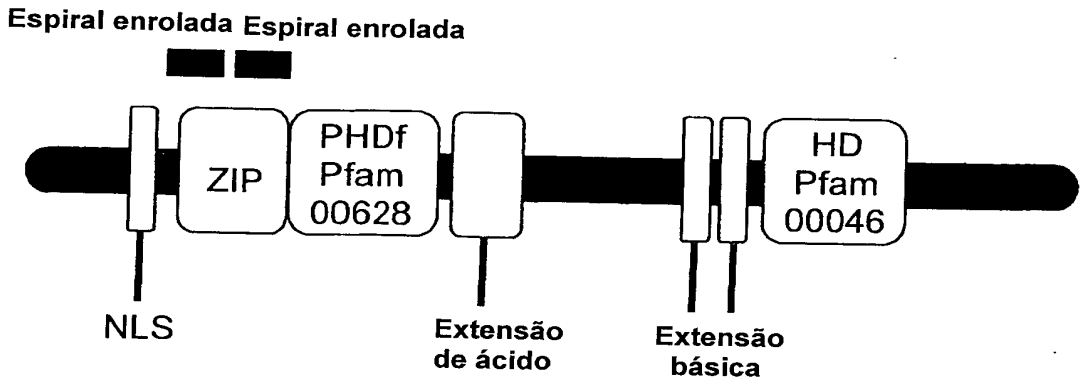


FIGURA 14



FIGURA 15

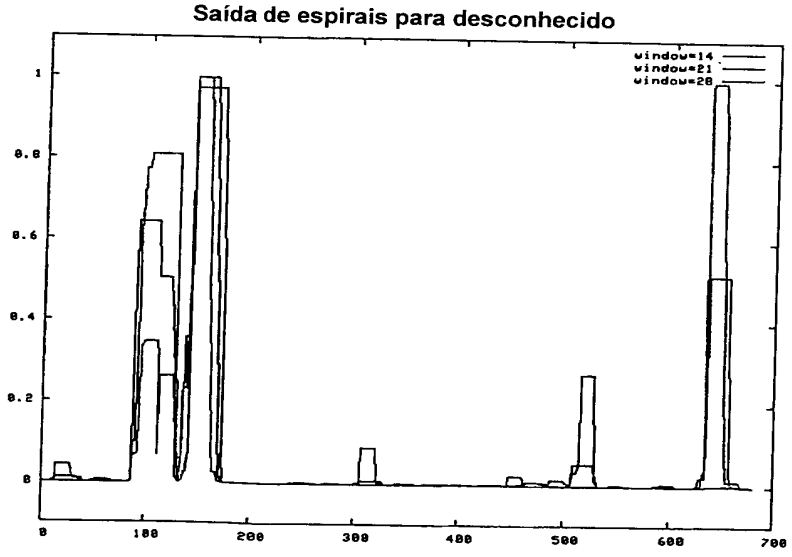


FIGURA 16

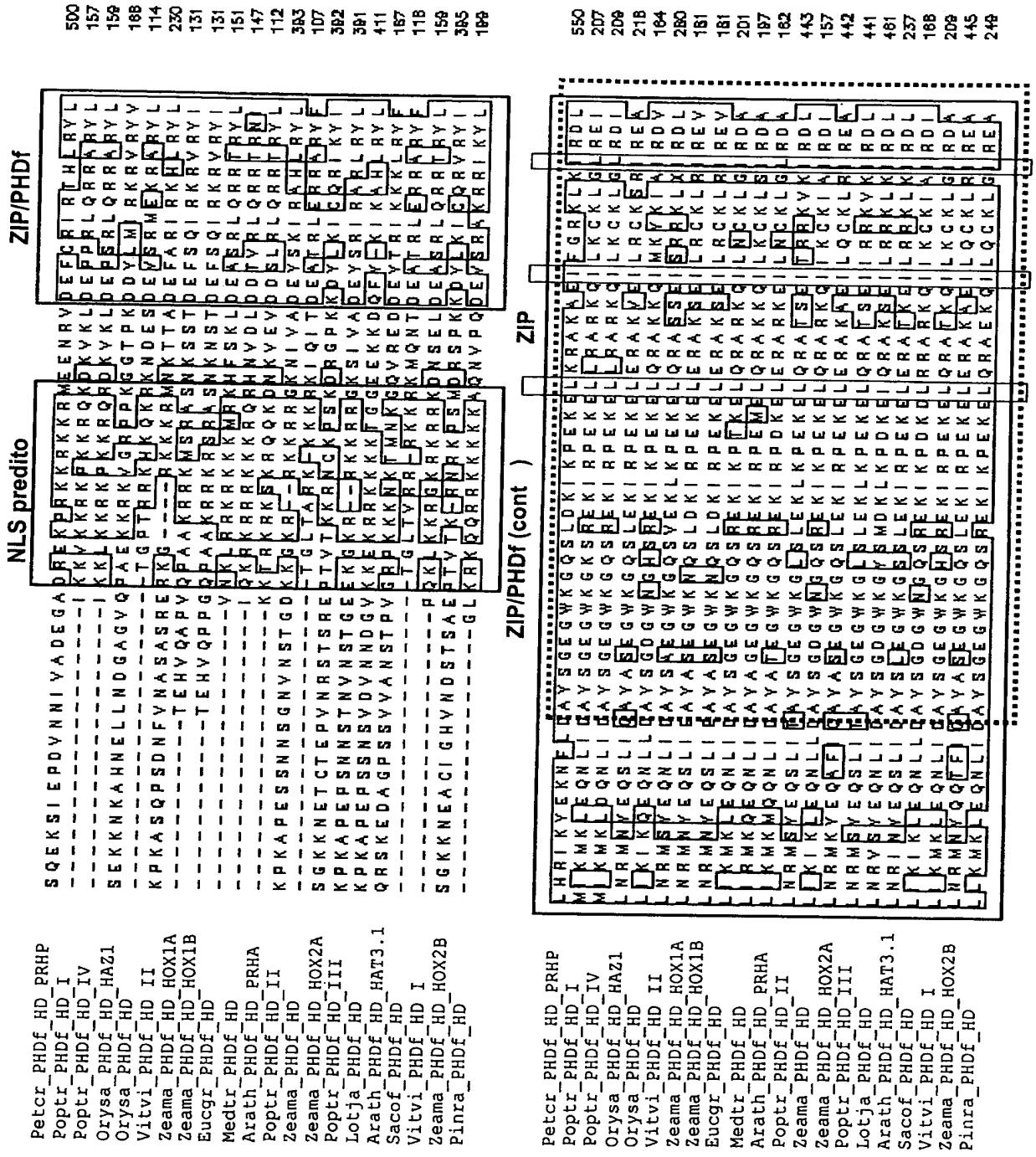

NLS predito
ZIP/PHDf
ZIP/PHDf (cont)
ZIP
Petcr_PHDf_HD_PRHP
Poptr_PHDf_HD_I
Poptr_PHDf_HD_IV
Orysa_PHDf_HD_HAZ1
Orysa_PHDf_HD
Vitvi_PHDf_HD_II
Zeama_PHDf_HD_HOX1A
Zeama_PHDf_HD_HOX1B
Eucgr_PHDf_HD
Medtr_PHDf_HD
Arath_PHDf_HD_PRHA
Poptr_PHDf_HD_II
Zeama_PHDf_HD
Zeama_PHDf_HD_HOX2A
Poptr_PHDf_HD_III
Lotja_PHDf_HD
Arath_PHDf_HD_HAT3.1
Sacof_PHDf_HD
Vitvi_PHDf_HD_I
Zeama_PHDf_HD_HOX2B
Pinra_PHDf_HD

FIGURA 17

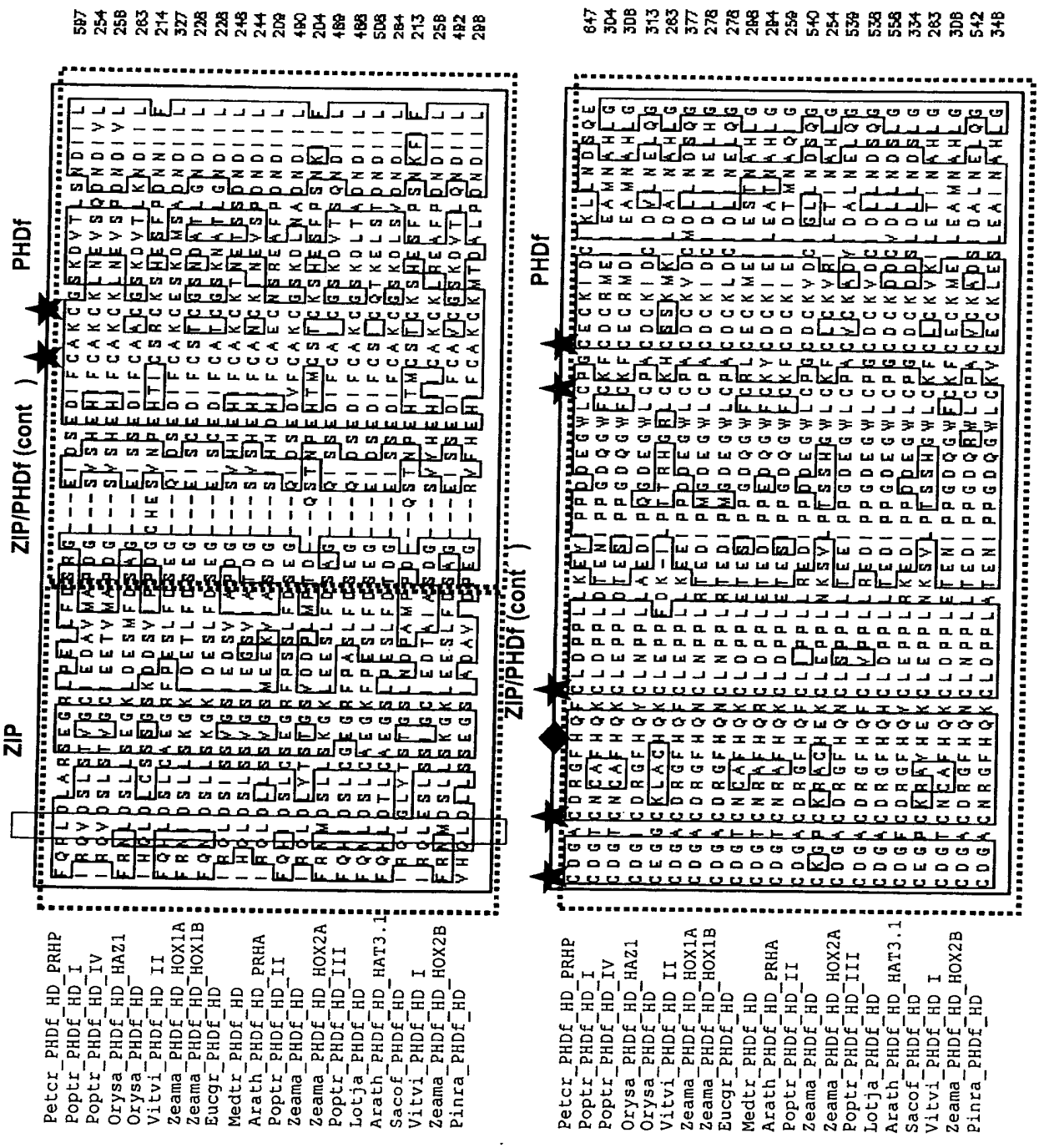

ZIP
ZIP/PHDf (cont)
PHDf
ZIP/PHDf (cont)
PHDf
Petcr_PHDf_HD_PRHP
Poptr_PHDf_HD_I
Poptr_PHDf_HD_IV
Orysa_PHDf_HD_HAZ1
Orysa_PHDf_HD_
Vitvi_PHDf_HD_II
Zeama_PHDf_HD_HOX1A
Zeama_PHDf_HD_HOX1B
Eucgr_PHDf_HD_
Medtr_PHDf_HD_
Arath_PHDf_HD_PRHA
Poptr_PHDf_HD_II
Zeama_PHDf_HD_
Zeama_PHDf_HD_HOX2A
Poptr_PHDf_HD_III
Lotja_PHDf_HD_
Arath_PHDf_HD_HAT3.1
Sacof_PHDf_HD_
Vitvi_PHDf_HD_I
Zeama_PHDf_HD_HOX2B
Pinra_PHDf_HD_

FIGURA 17

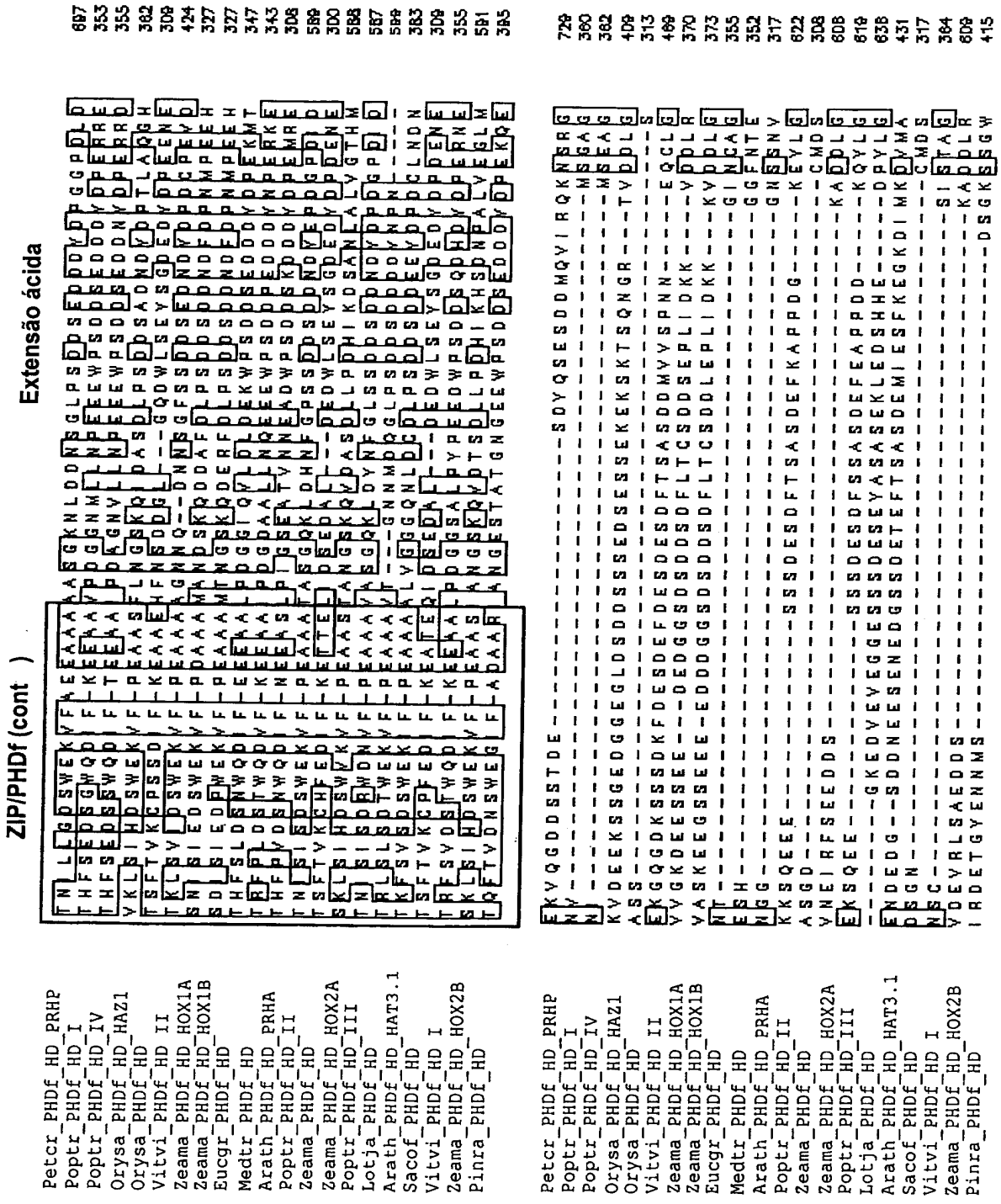

ZIP/PHDf (cont)
Extensão ácida
Petcr_PHDf_HD_PRHP
Poptr_PHDf_HD_I
Poptr_PHDf_HD_IV
Orysa_PHDf_HD_HAZ1
Orysa_PHDf_HD
Vitvi_PHDf_HD II
Zeama_PHDf_HD_HOX1A
Zeama_PHDf_HD_HOX1B
Eucgr_PHDf_HD
Medtr_PHDf_HD
Arath_PHDf_HD_PRHA
Poptr_PHDf_HD_II
Zeama_PHDf_HD
Zeama_PHDf_HD_HOX2A
Poptr_PHDf_HD_III
Lotja_PHDf_HD
Arath_PHDf_HD_HAT3.1
Sacof_PHDf_HD
Vitvi_PHDf_HD I
Zeama_PHDf_HD_HOX2B
Pinra_PHDf_HD

FIGURA 17

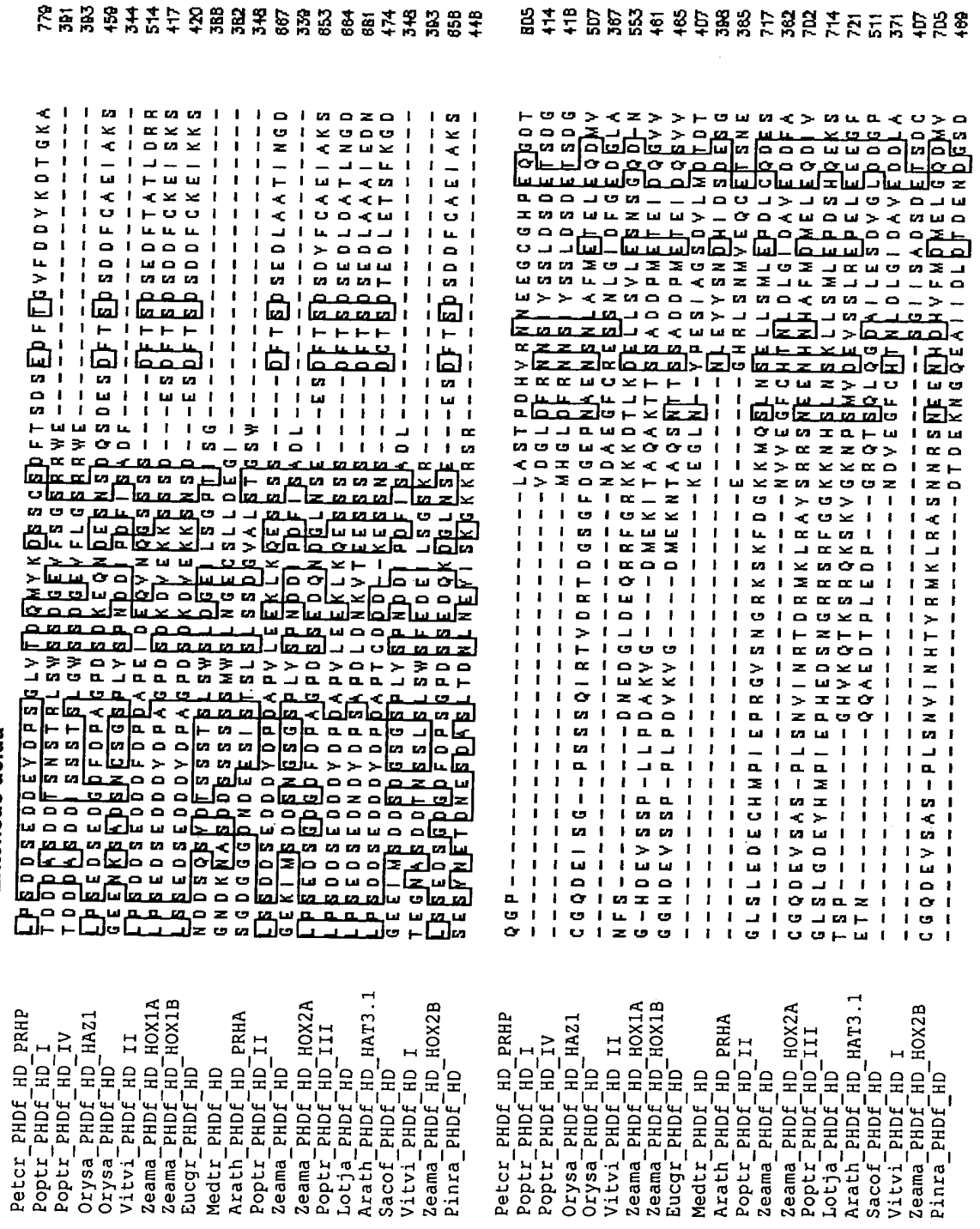


FIGURA 17


Extensão básica
Extensão básica
Petcr_PHDf_HD_PRHP
Poptr_PHDf_HD_I
Poptr_PHDf_HD_IV
Orysa_PHDf_HD_HAZ1
Orysa_PHDf_HD_
Vitvi_PHDf_HD_II
Zeama_PHDf_HD_HOX1A
Zeama_PHDf_HD_HOX1B
Eucgr_PHDf_HD_
Medtr_PHDf_HD_
Arath_PHDf_HD_PRHA
Poptr_PHDf_HD_II
Zeama_PHDf_HD_
Zeama_PHDf_HD_HOX2A
Poptr_PHDf_HD_III
Lotja_PHDf_HD_
Arath_PHDf_HD_HAT3.1
Sacof_PHDf_HD_
Vitvi_PHDf_HD_I
Zeama_PHDf_HD_HOX2B
Pinra_PHDf_HD_
Petcr_PHDf_HD_PRHP
Poptr_PHDf_HD_I
Poptr_PHDf_HD_IV
Orysa_PHDf_HD_HAZ1
Orysa_PHDf_HD_
Vitvi_PHDf_HD_II
Zeama_PHDf_HD_HOX1A
Zeama_PHDf_HD_HOX1B
Eucgr_PHDf_HD_
Medtr_PHDf_HD_
Arath_PHDf_HD_PRHA
Poptr_PHDf_HD_II
Zeama_PHDf_HD_
Zeama_PHDf_HD_HOX2A
Poptr_PHDf_HD_III
Lotja_PHDf_HD_
Arath_PHDf_HD_HAT3.1
Sacof_PHDf_HD_
Vitvi_PHDf_HD_I
Zeama_PHDf_HD_HOX2B
Pinra_PHDf_HD_

FIGURA 17

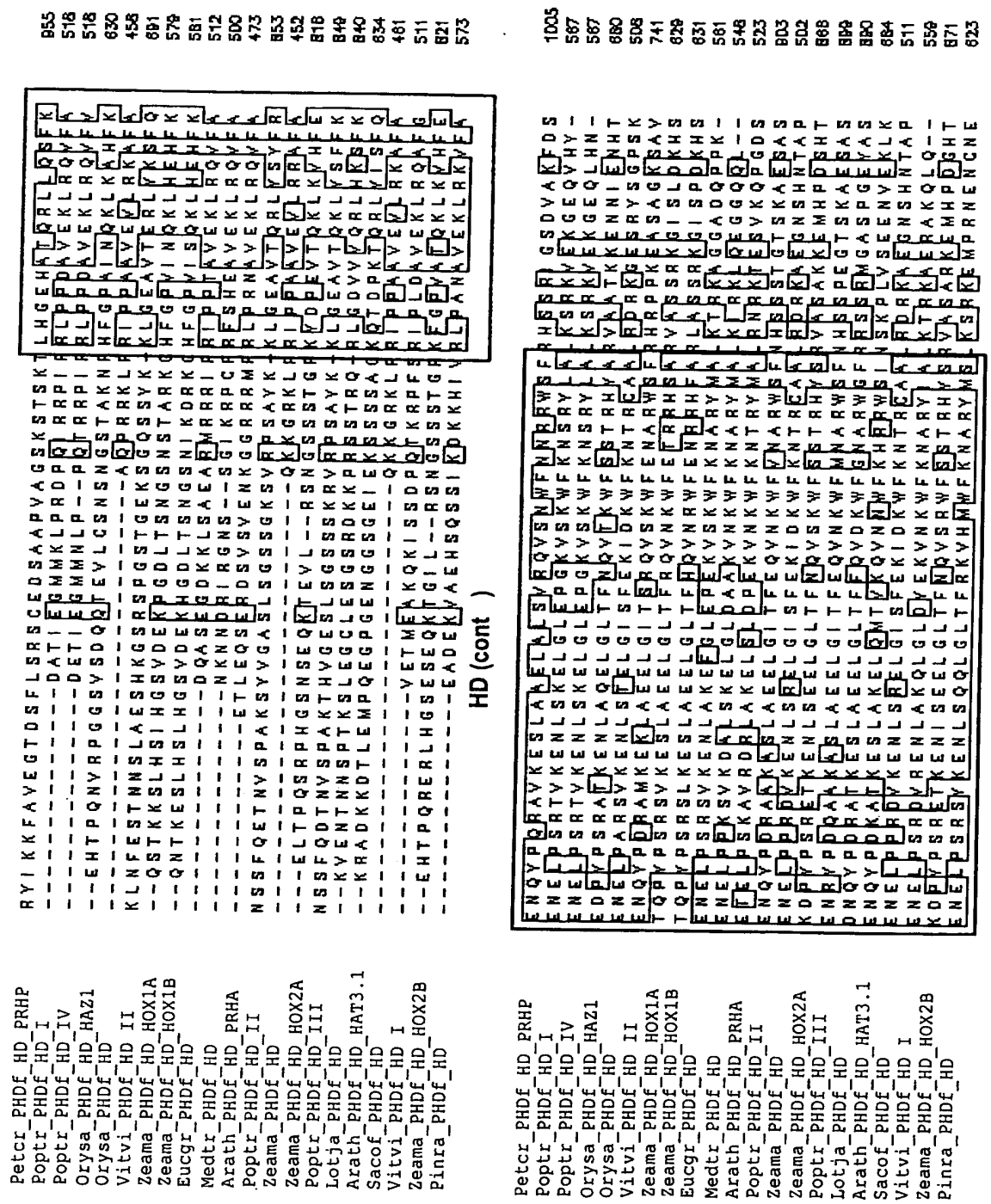

HD
HD (cont)
Petcr_PHDf_HD_PRHP
Poptr_PHDf_HD_I
Poptr_PHDf_HD_IV
Orysa_PHDf_HD_HAZ1
Orysa_PHDf_HD
Vitvi_PHDf_HD_II
Zeama_PHDf_HD_HOX1A
Zeama_PHDf_HD_HOX1B
Eucgr_PHDf_HD
Medtr_PHDf_HD
Arath_PHDf_HD_PRHA
Poptr_PHDf_HD_II
Zeama_PHDf_HD
Zeama_PHDf_HD_HOX2A
Poptr_PHDf_HD_III
Lotja_PHDf_HD
Arath_PHDf_HD_HAT3.1
Sacof_PHDf_HD
Vitvi_PHDf_HD_I
Zeama_PHDf_HD_HOX2B
Pinra_PHDf_HD

FIGURA 17

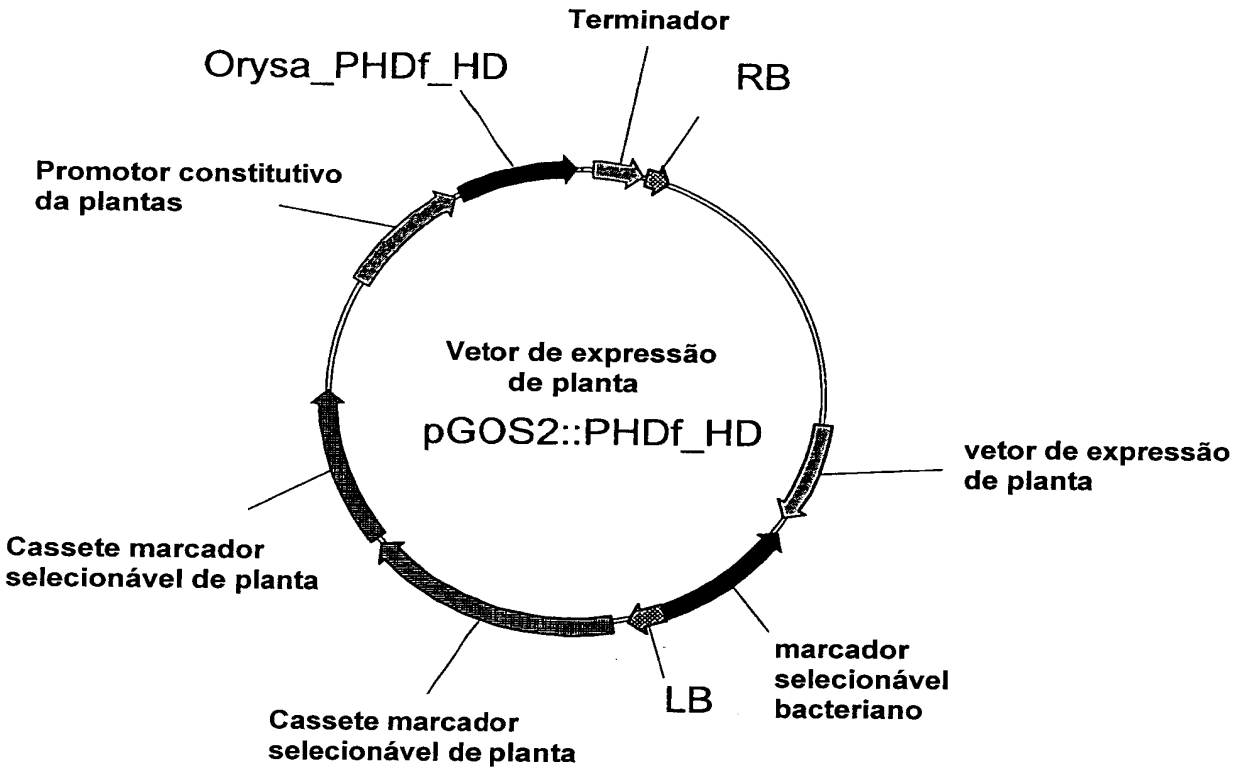

Terminador
Orysa_PHDf_HD
RB
Promotor constitutivo
da plantas
Vetor de expressão
de planta
pGOS2::PHDf_HD
vetor de expressão
de planta
Cassete marcador
selecionável de planta
marcador
selecionável
bacteriano
LB
Cassete marcador
selecionável de planta

FIGURA 18

SEQ ID NO: 179, sequência de ácido nucléico de PHfD_HD de Oryza sativa Os02g0147800 NM_001052422.1

ATGAATACCCCAGAAAAGAAACCACTGTGCTATACTTCTAGAAGAGCACTGCAACAGAGGACAGAA
TCCTCTTCGGAATTGATCTCTGTATCAAAAAGAGCAACCCGCCAGAATACTCCACGCAAGCCGGAT
TCTCCTCCCAAACGAACAACAAGAAGTAGCGCTAACCTGGCGAAATGTATAGAGAACAAACACCAT
AGCAGTCCTCTGAAGCGGCGGCGGGGTTCAGATGCGGCTACCGGAAAGTCTGCAACAGGGCCAACA
AGGAGGAAGCATAAACAGAAAAGGAAAAATGATGAGAGTGATGAGGTCAGTCGCATGGAAAAAAGA
GCAAGATATTTACTTATCAAAATAAAACAGGAGCAGAATTTGCTAGATGCATACTCTGGAGATGGT
TGGAATGGGCACAGTCGAGAGAAAATAAAGCCAGAGAAGGAGCTACAGCGTGCTAAGAAACAGATC
ATGAAATACAAAATCGCTATCCGTGACGTCATTCACCAACTTGATTTGTGTAGTTCCAGTGGGAGC
AAGGATGACTCTGTGATTCCACCAGATGGGTGTCACGAGTCTGTCAATCCTGAACATACAATATGT
TCAAGGTGCAAATCCCACGAGTCATTTCCTGACAATAACATTATATTCTGTGAGGGGGGCTGCAAA
CTGGCGTGTCATCAGAAATGTTTGGAGCCTCCTTTTGATAAAATTCTTCCAACTACCCGCCATGGG
CGGCTTTGCAAACACTGTTCTTCCAAGATGAAAATTTTAGACGCCATTAATGCCCATCTGGGGACA
AGTTTTACAGTGAAGTGTCCCTCCAGTGATATTTTCAAGGAAGCAGCTGAACATTTCAACTCTGAT
GATGGACTGGGCCAAGATTGGCTATCTGAGTATTCAGGTGATGAGGACTATGACCCTGAAGAAAAT
GAGGCCAGCAGCAGTGGTGAAGAGAATAAATCTGCCGATTCCAATTGTTCAGGGAGTCCACTTTAT
TCTCCAAATGATGATATTCCAGACTTCATATCAGCAGATTTCAATGATGCTGAAGGATTCTGTCGT
GAAAGTTCAAATCTAGGAATCGATTTTGGTGAAGATGGTTTGGCTGAGATTCTTACCCACCAAAGG
CCAAGAAGAGATGTTGACTATACACAACTTAATGAGCAAATGTTTGGAGAGCCTATTGGCAATGAC
GAACAGAGTGAAGATGAAGATTGGGGTCTTAACAAGAGAAAGAAAAGAAGAACTGGTTCAACTGGT
GTTGGGACTAATTCCGTAGAGGGTCGATCAGATGTCAAATCCAATAAGAAAGCCCAACCTCGGAGA
AAACTTTTCAGGATTCCTCCTGCGGCAGTTGAGGTGCTCCGTAAAGCTTTTGCTGAAAATGAGCTT
CCTGCCCGGAGTGTCAAAGAAAATCTCTCAACAGAGCTGGGTATTTCTTTTGAAAAGATTGATAAG
TGGTTCAAAAATACACGGTGTGCTGCTCTCAGGGATAGGAAGGGTGAAAGTCGTTATTCTGGTCCC
AGCAAAAGGTCAAGAACAAGCATAGAAAAGGCTGAAACTTCAGCTAAAGTGGATCAAATGGACAAC
TCATGCTTTCTTCCCTTATCTGAAATAATCAACGTGCCCACACGTCTGCAGAAAGGTCTTGATAAG
AAGCCAAAATCAATTAACAGCCCTCCAAGACCTCAGGACAATGAAACTTGCTTGTCGCCAACTGAT
AAAACCAAGGAGGGTACACCGCCTACGATCAAGCCTTCCATCACAGATTCTAGTCAACTGATGAAT
AACGACATCGGCACTGAGGAGACTGCTGTTTCTTGGGTGGACACTTGGGCCTCTGACGCTCTGCAT
TTCTTGGACGTGAGCGATGATGAACATTTCTTCGATGTGATCGAGAAGGTGTGCGGTCTCGAGAAT
CGGCTGCAGCGGTTGAAGGAGAACATGCTATCATCGTCATCGTCAACTGACAACAATGTGGCTGCT
GAGAGTGGCTTGCAAAACGAAGTTGTGCTTGTACCAGCTGCTGAGCTCAAGGATAAAGCATCTTAA

SEQ ID NO: 180, sequência de polipeptídeo traduzida de PHDf_HD Oryza sativa

MNTPEKKPLCYTSRRALQQRTESSSELISVSKRATRQNTPRKPDSPPKRTTRSSANLAKCIENKHH
SSPLKRRRGSDAATGKSATGPTRRKHKQKRKNDESDEVSRMEKRARYLLIKIKQEQNLLDAYSGDG
WNGHSREKIKPEKELQRAKKQIMKYKIAIRDVIHQLDLCSSSGSKDDSVIPPDGCHESVNPEHTIC
SRCKSHESFPDNNIIFCEGGCKLACHQKCLEPPFDKILPTTRHGRLCKHCSSKMKILDAINAHLGT
SFTVKCPSSDIFKEAAEHFNSDDGLGQDWLSEYSGDEDYDPEENEASSSGEENKSADSNCSGSPLY
SPNDDIPDFISADFNDAEGFCRESSNLGIDFGEDGLAEILTHQRPRRDVDYTQLNEQMFGEPIGND
EQSEDEDWGLNKRKKRRTGSTGVGTNSVEGRSDVKSNKKAQPRRKLFRIPPAAVEVLRKAFAENEL
PARSVKENLSTELGISFEKIDKWFKNTRCAALRDRKGESRYSGPSKRSRTSIEKAETSAKVDQMDN
SCFLPLSEIINVPTRLQKGLDKKPKSINSPPRPQDNETCLSPTDKTKEGTPPTIKPSITDSSQLMN
NDIGTEETAVSWVDTWASDALHFLDVSDDEHFFDVIEKVCGLENRLQRLKENMLSSSSSTDNNVAA
ESGLQNEVVLVPAAELKDKAS

FIGURA 19

SEQ ID NO: 181, sequência de ácido nucléico de PHDf_HD_PRHA de Arabidopsis thaliana NM_119140

ATGGAGGAGAGTGAAACAAAGGGGAGAATCAGTCAAGAAACTGATAAGGCTTGTGTTTCAGTTGAG
AGAATTGGGTCTACTTTACTCTCATCATTTGTGAAGAAAGGTAAAGAAGTATCAAATAAGAGAAAT
TCTAAGCAGAATAAGCGAAAAGCGGAAGAAGAACTCTGTTCGAAATCAAGAACCAAGAAGTACTCT
CGGGGATGGGTCCGTTGTGAGGAAATGGAGGAGGAGAAGGTTAAGAAAACAAGAAAGAGGAAAAGT
AAGAGACAGCAAAAGGATAATAAGGTTGAGGTTGATGATTCTTTAAGGTTGCAAAGGAGGACAAGG
TATTTGCTCATTAAGATGAAGATGCAGCAGAATCTTATCGATGCATACGCAACTGAAGGTTGGAAA
GGTCAGAGCCGGGAAAAGATAAGACCAGACAAGGAGCTTGAGAGAGCCCGGAAAGAGATCTTAAAC
TGCAAGCTTGGACTACGCGACGCCATTCGCCAACTGGATCTGCTTAGTTCTGTGGGAAGCATGGAA
GAGAAAGTGATAGCTTCAGATGGATCTATTCATCATGATCATATATTTTGTGCAGAATGCAATTCT
CGAGAAGCTTTTCCAGACAATGATATAATACTTTGTGATGGAACGTGTAATCGAGCATTTCACCAG
AAATGCCTTGACCCTCCTTTGGAAACAGAAAGCATACCTCCCGGAGATCAGGGCTGGTTTTGCAAA
TTTTGTGATTGCAAAATTGAAATTATCGACACAATGAATGCACAAATTGGAACTCATTTCCCTGTG
GACAGCAACTGGCAGGATATCTTCAATGAAGAAGCTAGTCTTCCTATTGGATCTGAAGCTACAGTT
AACAATGAAGCAGATTGGCCTTCGGATGACTCCAAGGATGATGATTATGACCCTGAAATGAGGGAA
AACGGTGGAGGAAACAGCAGCAACGTTAGTGGTGATGGTGGTGGTGATAACGATGAAGAAAGCATT
TCAACTAGCTTGAGCTTATCATCTGATGGTGTTGCCCTTTCAACGGGATCATGGGAAGGCCATAGA
CTCAGCAATATGGTGGAGCAGTGCGAAACAAGTAATGAAGAAACTGTATGTGGGCCAAGACAGCGA
AGGACTGTTGACTATACACAGTTGTACTATGAAATGTTTGGAAAGGATGCAGTTTTGCAAGAGCAA
GGTAGTGAAGATGAAGACTGGGGTCCGAATGACAGAAGGAAGAGAAAAAGGGAATCTGATGCAGGA
AGCACACTTGTAACTATGTGTGAAAGCAGTAAGAAAGATCAAGATGTTGTAGAAACGCTAGAACAG
AGTGAGAGAGATTCTGTCTCTGTGGAAAATAAAGGAGGTCGAAGGCGGATGTTCAGACTCCCAAGA
AATGCAGTTGAGAAGCTACGTCAAGTATTTGCAGAGACTGAGCTTCCCTCAAAAGCTGTGAGGGAT
CGTCTTGCAAAAGAGCTGAGTCTTGATCCAGAGAAGGTCAATAAATGGTTTAAAAATACTCGGTAT
ATGGCACTGCGGAATAGGAAGACTGAGAGCGTGAAACAACCTGGGGATTCTAAGACAGTCTCTGGA
GGAGATTCTGGACCAGAAGCAGTCATGGAGAACAACACTGAAACAAATGAAGTTCAAGATACTTTG
GATGACACTGTTCCACCCGGATTTGACGCTACAAATCAAAACATACTCTCCCCATGTAATAACAAC
CAGGAGGAATTTCAACAGGAAAATGTCTCTTTTCCTTCACCCACAGATGAAAGTCAACAGTATCTG
GAACAAAACGATTCTTCTTTCGTTCTAGTGCCACATGAAAAGCAAAGCAGCGAAATTAGCTTAAAG
ACGGCTGTAGAGGAGAATGAGACAGAAAGCAAGATGATGAAAGAGCCACATGAAGAGCTAAGCAGT
GAAATGAGCCTAAAGACGGCTGCTGAGGAGAAGGAGACAGAGAGCAAGATGATAGAAGAGCCACAT
GAAGAGCTAAGCAGAGAAATGAGCTTAAAGACGGCTGTAGAGGAGAAGGAGACAGAGAGCAAGATG
ATGGAAGAGCCACATGATGAGCTAAACAGCGAAATGAGCTTAAGTACAGCGGTAGAGGAGAAGGAG
ACAGGGAGCAAGATGACGGAAGAGTCACATGAAGAGCTAAGCAACGAAATGAGCTTAGAAGAGAAG
GAGACGGGGAGAAAGATGACGGAAGAGGAAGAATTGGAAGCAGTAATGGAGATGCTTTGTAGAACA
GAAAACAAGTTACTGGACGTGACGCAGAGGCTTGATAGATTTAAAACACCAAAAGGCCGAAAAAAG
TTAGGCAATTCTTCTTCTCCTTTACTTGAAGAAGACTCAGTAGTGTATGTTCCAATAGCAGAGATC
AAGGAGAAAAGGTAA

SEQ ID NO: 182, sequência de polipeptídeo de PHDf_HD_PRHA de Arabidopsis thaliana

MEESETKGRISQETDKACVSVERIGSTLLSSFVKKGKEVSNKRNSKQNKRKAEEELCSKSRTKKYS
RGWVRCEEMEEEKVKKTRKRKSKRQQKDNKVEVDDSLRLQRRTRYLLIKMKMQQNLIDAYATEGWK
GQSREKIRPDKELERARKEILNCKLGLRDAIRQLDLLSSVGSMEEKVIASDGSIHHDHIFCAECNS
REAFPDNDIILCDGTCNRAFHQKCLDPPLETESIPPGDQGWFCKFCDCKIEIIDTMNAQIGTHFPV
DSNWQDIFNEEASLPIGSEATVNNEADWPSDDSKDDDYDPEMRENGGGNSSNVSGDGGGDNDEESI
STSLSLSSDGVALSTGSWEGHRLSNMVEQCETSNEETVCGPRQRRTVDYTQLYYEMFGKDAVLQEQ

FIGURA 19 CONT.

GSEDEDWGPNDRRKRKRESDAGSTLVTMCESSKKDQDVVETLEQSERDSVSVENKGGRRRMFRLPR
NAVEKLRQVFAETELPSKAVRDRLAKELSLDPEKVNKWFKNTRYMALRNRKTESVKQPGDSKTVSG
GDSGPEAVMENNTETNEVQDTLDDTVPPGFDATNQNILSPCNNNQEEFQQENVSFPSPTDESQQYL
EQNDSSFVLVPHEKQSSEISLKTAVEENETESKMMKEPHEELSSEMSLKTAAEEKETESKMIEEPH
EELSREMSLKTAVEEKETESKMMEEPHDELNSEMSLSTAVEEKETGSKMTEESHEELSNEMSLEEK
ETGRKMTEEEELEAVMEMLCRTENKLLDVTQRLDRFKTPKGRKKLGNSSSPLLEEDSVVYVPIAEI
KEKR

SEQ ID NO: 183, sequência de ácido nucléico de PHDf_HD de Eucalyptus grandis ADW17964

ATGCATTCCAGAAAGAATCTTGCCCAAAAAGAAGACTTTTTGAAGGCAGAATCTGGATCGGAGGTT
CTTGCCTCGCTGATATTGAGAAATAGGAGAAGGAAAAAGCACAAACAGAAACTACAATCTTCTGTA
AAAGCAAGAGGCTCAACCCTTGCCAAGAAGAGGGCCAATAACTCCATCAGAAAGAGACTGGTGCGT
AGGAACAGCTCTGATAAAGACAAACTCATAAATTTAAAAGCTTTTCATAATAAGAAAATTGAAATA
GTTCCCGAGAAGCTTTCATCCCCTCCATCTTCAGAAAGAGAGCTGCCTATTGCTTCTGACGAAAAT
GTGCAAAACACCGATGAGGATGTCAAAGTTAACAAGCTGAGAAGGAAGAGGAAGAAGAAAATGAGA
AAGCATTTTTCGAAGTTAGATGAGGCTTCACGATTGCAGAGGAGAACAAGATACCTTCTTATTAAA
ATGAAGCTGGAACAGAACCTTATTGATGCTTACTCTGGTGAAGGCTGGAAAGGCCAGAGTCGGGAA
AAGATCAAGCCAACAAAGGAATTACAAAGGGCCAGGAAGCAGATACTGAATTGCAAACTCGGAATA
CGTGATGCTATTCGACAGCTAGACTCCATTAGTTCAGTGGGCAGCATTGAAGACTCTGTGATTGCT
CCTGATGGTTCTGTTCACCATGAACATATATTCTGTGCAAAATGCAAGACAAATGAAACTTCCTCA
GATAATGATATCATACTCTGCGATGGGACATGCAATTGTGCTTTTCATCAAAAATGTCTTGATCCT
CCACTGGAAACTGAAAGCATTCCCCCAGGAGATCAGGGCTGGTTTTGCAGACTCTGCGAGTGTAAG
ATGGAAATTATAGAGTCAACAAATGCACATCTGGGGACCCACTTTTCTCTTGACAGCAACTGGCAG
GATATATTCGAGGAAGAAGCTGCATTGCCTGATGGGGGAATTCAGGTGCTTGACCTAGACGAAAAG
TGGCCTTCAGATGACTCTGAAGATGATGACTATGACCCAGAAAAGATGACGAACACCGGCATCAAC
TGTGCCGGCAATGATGATTCTCAATCATACGACACAAGTAGCTCTACCAGCCTGAGTTGGTCATTA
GATGGTGAAATTTTGTCAGGACCAACAATATCAGGCAAGGAAGGGTTGAACTATCCTGAGAGCATT
GCAGGGTCTGATGTATTAATGGACACGGATACTGATACTGTATGTGGCCCACGACAGCGAAATGCT
GTTGATTATAAGCAGCTTTATAATGAAATGTTCGGTAAGGATACTCCTGCCTGTGAACAAGCCAGT
GAAGATGAAGACTGGGGTCCTAATAAAAGAAGAAGAAGAGAAAGGGAAACAGATGCAGCCAGCACC
CTCATGACTCTCTGTGAAAGTGAGAAAAACAATCCATTCATTGACCAAGCCTCAGAAGGTGATAAG
AAACTCTCTGCAGAAGCAAGAATGAGAAGGCGGATTTTCAGAATTCCTCCAACTGCAGTTGAGAAG
CTTCGCCAAGTATTTGCAGAGAATGAACTTCCCTCAAGATCTGTCAAAGAAAATCTTGCTAAAGAG
TTTGGTCTTGAACCTGAGAAGGTTAGCAAATGGTTCAAGAATGCACGATACATGGCACTCAAAACA
AGAAAGGCTGGGGGGGCTGATCAACCTCCAAAATCCCCGCAAAAGCATAAAGAATCTGAACTAGCA
AAAATGATGAAAGGGGCTGCGCCTATTTTATCTAAGGATGCCTCAACAGACATTGAATTCCATTCC
TCTCTAAATACTGAGAAAGTCATCCAGCAAAAGAGTACGAAGCCAATACACGGGAAGGGTTCACCT
ATGGTACCTGAACATAGCTATATGGAGAATTTTGAGTTCGATGACAATATGTCCTTGAAACAAATA
AAGCTATTGAAAGGAAAACCAAAGGACAGGAAGAGGGTTGACTTCGCAGCCAGTGAGGGATTACGA
GAAAGAGAAGCTCAGTTGCGGGAGCTATGTAAAATTAGAAGTAGGCTAGATAGTCTGAGGAAGAGA
ACCCTTACATTTCAGTCGAGTGAATCCAAGGCTTCAGACGAATTCAACACTTGTAGATCACGTGTA
GTATATGTTCCAGTAGCAGAAATAAAGGAGAGAGCTTGA

FIGURA 19 CONT.

SEQ ID NO: 184, sequência de polipeptídeo traduzida de PHDf_HD de Eucalyptus grandis

MHSRKNLAQKEDFLKAESGSEVLASLILRNRRRKKHKQKLQSSVKARGSTLAKKRANNSIRKRLVR
RNSSDKDKLINLKAFHNKKIEIVPEKLSSPPSSERELPIASDENVQNTDEDVKVNKLRRKRKKKMR
KHFSKLDEASRLQRRTRYLLIKMKLEQNLIDAYSGEGWKGQSREKIKPTKELQRARKQILNCKLGI
RDAIRQLDSISSVGSIEDSVIAPDGSVHHEHIFCAKCKTNETSSDNDIILCDGTCNCAFHQKCLDP
PLETESIPPGDQGWFCRLCECKMEIIESTNAHLGTHFSLDSNWQDIFEEEAALPDGGIQVLDLDEK
WPSDDSEDDDYDPEKMTNTGINCAGNDDSQSYDTSSSTSLSWSLDGEILSGPTISGKEGLNYPESI
AGSDVLMDTDTDTVCGPRQRNAVDYKQLYNEMFGKDTPACEQASEDEDWGPNKRRRRERETDAAST
LMTLCESEKNNPFIDQASEGDKKLSAEARMRRRIFRIPPTAVEKLRQVFAENELPSRSVKENLAKE
FGLEPEKVSKWFKNARYMALKTRKAGGADQPPKSPQKHKESELAKMMKGAAPILSKDASTDIEFHS
SLNTEKVIQQKSTKPIHGKGSPMVPEHSYMENFEFDDNMSLKQIKLLKGKPKDRKRVDFAASEGLR
EREAQLRELCKIRSRLDSLRKRTLTFQSSESKASDEFNTCRSRVVYVPVAEIKERA

SEQ ID NO: 185, sequência de ácido nucléico de PHDf_HD de Medicago truncatula AC123547

ATGCGGGATACTGAGAGGTTAAATAACCAAGGATCTACAAAACCGAGTAACACAAAGGAACATATC
AAGTTGAAGGTAGATTCCCCACAGTTCAAAACAAGTAGCACAAAACACCAAAGGAAGAAACACAGA
CTGAAATCAAAATCTCATAAACTCGGAGGTTGCACAAATGCGTCAGTGAGGACAGCTACTGATTCT
TCCAACAAGGCGTCAACCAAGGATTCTTCAAATAAAACGGATAGGAATTCAACACAGGTTCAGCCT
TCAAAAAAGATACAAGGAGGAAAAACTTCTCTTAATAATGATAGGAAAGGTGAGAAAGATGTTGTT
GATCAAGAGGGAAATATTCAAAAACGTAAGAGAAGGAGAAAGAAGAAAAGACAAAGACATAATGTA
GATCTTGATGATACTGTACGCCTCCAGAGAAGAACAAGGAATATCTTAATCAGAATGAAGCAAGAA
CAGAATCTTATCGATGCTTACGCTGGAGAAGGTTGGAAAGGTCAGAGTCGAGAGAAGATTAGGCCA
GAAATGGAGCTACAAAGAGCAAAAAAGCAGATTTTAAAATGCAAGCTTAGTATCCGTGATGCCATT
CACCAACTGGATTCTCTTAGTTCAGTTGGTAGTATTGAAGGTTCTGTCATTGCTACAGATGGATCT
GTTTCTCATGAACATATATTCTGTGCAAATTGCAAGATAAACGAAGTTTCTCCAGATAATGATATT
ATACTCTGTGATGGCACATGCAATCGTGCGTTTCACCAGAGATGTCTTGATCCTCCTTTGGAGACT
GAAGACATTCCTCCTGAAGACCAAGGCTGGTTTTGCAAGTATTGTGACTGTAAGATAGAAATATTG
GAGGCAACAAACGCGCATCTTGGGACTCGTTTCCCCTTAGACAGTACTTGGCAGGATGTATTCAAG
GAAGAAGCTGCTATTCCTGACGGCGATGCTGCATTACTAAATCAAGAAGAAGAATGGCCGTCAGAT
GATCCTGAAGACGATGACTATAATCCAGAGAGGAAAGAAGAAAGCCACGGTGGCTTCAACACAGAA
GGAAATGATAAAAACGCATCTGATGATTCAAGCAGTTCTTCCAGCATGTGGTCTCTAAATGGAGAA
TGCTCTCTGTTAGATGAAGGCATCAACCTTGAATATTATTCTAACGACCACATAGATTCTGATGAA
TCCGGGGAAATAGCATGCGGTCGTAGGCAGCGTAGAGCTGTTGACTACAAGAAACTTTATGATGAA
ATGTTTGGAAAGGATGCTCCACCTTGTGAACAAGTGAGTGAAGATGAAGATTGGGGTCCTCGAAAA
AGACGGCGAAGAGAAAAGGAGTCTGAGGCTGTTAATACTCTTATGACTCTACATGAGAGTGAAAAT
AAATATCCGAATAATAAGAACAATGATAGAATAAGAGGGAATTCTTCAGGCATAAAAAGACCTTGT
TTTAGGTTTTCACATGAGGCAGTTGAGAAGCTTCGCCAAGTTTTTGCAGAGAACGAGCTTCCTCCA
AAATCCGTGAAGGATGCCCTTTCAAAAGAATTGGGACTTGATGCTGCAAAGGTTAACAAATGGTTC
AAAAATGCACGTTACATGGCACTTAAAATAAGAAAGCTTCAAGAAGGAGGACAACAGCTTCAGAGT
ATCACTTCTAAGACCTCAAAGGATTCTACATCACAACATGTGCAGGAAGATGAGGTTTTGAACCCG
AAGTCCGCGAAAATTACTGTGATTAGTTCCCTGAAGAAGTGTGAAAATGTTACTGGCAAAGAAAAA
ACAAAAGCATCTAGCAATCCTTTGAAGAAAAAGCGACCAGAAATACCTCCATGCTTGGGAGAAAAT
GGCAACAAGGATTCCATGGAGGTCAGTGATGATGTTAGCTTAATGAAACTATTACAAGACAGGAAA
AAGAAGGTAAGTTTTGCATTTGAAGGAGACTCTGAGGCAGCAGAGTTGGAATTTGAAAGACTCAGC
AAAGTGAAGACGAAAATAGATAGATTGAAGCAAAGATTAACTGGAGTTCAAAATTGCAGATCAAAG
GGTTCAGATGAAGTTCATTTGAATGAGCCATCTATTATGTATGTTCCTATAGCTGAGTTAAGGGAA
AAGGTGAAATAG

FIGURA 19 CONT.

SEQ ID NO: 186, Medicago truncatula sequência de polipeptídeo traduzida de PHDf_HD

MRDTERLNNQGSTKPSNTKEHIKLKVDSPQFKTSSTKHQRKKHRLKSKSHKLGGCTNASVRTATDS
SNKASTKDSSNKTDRNSTQVQPSKKIQGGKTSLNNDRKGEKDVVDQEGNIQKRKRRRKKKRQRHNV
DLDDTVRLQRRTRNILIRMKQEQNLIDAYAGEGWKGQSREKIRPEMELQRAKKQILKCKLSIRDAI
HQLDSLSSVGSIEGSVIATDGSVSHEHIFCANCKINEVSPDNDIILCDGTCNRAFHQRCLDPPLET
EDIPPEDQGWFCKYCDCKIEILEATNAHLGTRFPLDSTWQDVFKEEAAIPDGDAALLNQEEEWPSD
DPEDDDYNPERKEESHGGFNTEGNDKNASDDSSSSSSMWSLNGECSLLDEGINLEYYSNDHIDSDE
SGEIACGRRQRRAVDYKKLYDEMFGKDAPPCEQVSEDEDWGPRKRRRREKESEAVNTLMTLHESEN
KYPNNKNNDRIRGNSSGIKRPCFRFSHEAVEKLRQVFAENELPPKSVKDALSKELGLDAAKVNKWF
KNARYMALKIRKLQEGGQQLQSITSKTSKDSTSQHVQEDEVLNPKSAKITVISSLKKCENVTGKEK
TKASSNPLKKKRPEIPPCLGENGNKDSMEVSDDVSLMKLLQDRKKKVSFAFEGDSEAAELEFERLS
KVKTKIDRLKQRLTGVQNCRSKGSDEVHLNEPSIMYVPIAELREKVK

SEQ ID NO: 187, sequência de ácido nucléico de PHDf_HD de Pinus radiata ADW18458

ATGCCGAAAGCAGGAAAAGGCTCTTCCGGTTCACTCCGAAAGAAGCAAGCGTTGTCAAATACAGCA
TCAGAGCTAAGGCCTAAAGCTGCAATTCAGAAGACTAACAATGCCGGGAGTATTTCTGGCAGGATG
AGGGAAAGGAAACCACGGTTACATGTACAGATGATTGCCTCAATTCTCACGAAAAAAAGGACAGGT
GAAGCATCTGAATCGAGAAATGATTCTGAGCTAGAAGACAAACAATATCTGAGCAATAGTTCAAAA
ACAAATGCCAATTTTGTAGCGGGGAAGTTAGTTAAAAAGACTTTGAAAAAAGCCCCAAGATTAAAC
TTGATTAAGAATGACAGTCAGAAAAAACCCAGTAAGAGTTCTGCTTCGAGGTCTTATGCTGCACAG
AAACCTGAACGGCCACAACAGAGAAAAATTAGGAGGACAAAGTGGAAGGGTGATTCTAAAAATATA
ACAAAGGTGAAAACTGTGAGAACTAGTAATATAGATGGCTTGAAGAGGAAGCAGCAAAGGAGAAAG
AAGAAAAAGAAGGCGCAGAATGTTCCACAGGATGAAGTATCACGTGCAAAGAGAAGGATCAAATAT
CTTTTGCTCAAAATGAAACTTGAGCAGAATCTTATTGATGCATATTCAGGGGAGGGCTGGAAAGGC
CAGAGTCGGGAGAAAATAAAGCCAGAGAAGGAGTTACAGAGAGCTGAAAAGCAGATTCTGCAATGC
AAACTTGGAATTCGTGAAGCAGTGCATCAGCTTGATCTTCTAAGTTCGGAAGGATCTATAGCAGAT
GCCGTTATTGATCCAGAAGGGCGTGTATTTCATGAGCATATATTTTGTGCAAAATGCAAAATGACA
GATGCTCTTCCGGATAATGACATAATATTGTGTGATGGAGCTTGCAATCGTGGCTTCCATCAGAAA
TGTCTGGATCCACCTTTGGCTACTGAAAACATTCCTCCTGGAGATCAGGGTTGGCTCTGTAAAGTT
TGCGAATGTAAGCTAGAGAGTTTGGAAGCAATAAATGCTCATTTGGGCACCCAGTTTACAGTTGAT
AATAGCTGGGAGGGAATTTTTGCTGATGCAGCTAGAATTGCGAATGGTGAAAGTACAGCTACTGGG
AATGGGGAAGAATGGCCATCAGATGATTCAGAAGATGATGACTATGATCCAGAGAAACAAGAGATT
AGAGATGAAACGGGCTATGAGAATAATATGAGTGATTCTGGAAAGTCAAGTGGATGGTCAGAATCA
TATAATGAAACTGACAATGAATCAGATGCTTCCTTAACAGATAATCTAAATGAATACATCTCTAAG
GGAAAGAAGAGGAGCAGAGACACTGATGAAAAAAACGGACAAGAAGCTATAATAGATTTAGATACT
GATGAGAATGATGGCTCAGACATGCCAGTTTCAGGGAGAAGGCAGCGAAGAGATGTTGATTATAAA
AAGTTGCATGATGAAATGTTTGGAAAAGATGGACCTTGTGAAGATGAAGTCAGTGAAGATGAAGAC
TGGGGACCCAATAGGAGGCGTAGAAGAGCGAAACAGCCGGATACAAATACCACTAAGGATTCTTCA
AATGAAGAGAAAGAGCATGTTAGAAAGGAAGCAGATGAGAAAGTAGCAGAGCATTCACAAAGTAGC
ATAAAGGATAAGAAACATATTGTCAGACTGCCAGCTAATGCAGTTGAGAAGCTGCGCAAAGTCTTT
GCTGAAAATGAGCTACCGTCAAGGTCTTACAAGGAAAATCTGTCACAGCAACTGGGCCTTACATTT
CGAAAGGTCCATATGTGGTTCAAAAATGCCCGTTATATGTCTCTGAAGAGCCGAAAGGAAATGCCT
AGAAATGAGAACTGCAATGAAACTAATCTCATGGGAGAAGGTGATATGGGGCCAAAGAAAAAAATA
CCAAAAGAGAAGAAAGCAGAATCAACATGTCTTTTATCTTCTGTTAGAGCAGAATATAAATCTAGA
AATACAAAAAGGAAGCATAGGAGAAAAAACATCAAAGCACCCTTATCTATCTCTTCCAGAAGTGAG
AGAAAAAGAACTATGTGCAGACTGCCTGCCAATGCAGTTGAGAGGTTTCGCCAAGTCTTTGCTGAG

FIGURA 19 CONT.

AATGAGTTGCCTTCAATGTCCTGTAAAATGGCATTATCAAAGCAGTTGGATCTACCTTATCGGCAG
GTGCATATGTGGTTCAAAAATGCTCGTTATATGTCTTTGAAGAAGAAAAAGATGAACCGTTTAAAA
GAATCAGCTCCCGTTATTTCCTCACATACCAGAAGACTTGAAGTAGAAAGTAGCAAACCCAATCCA
CAAATCAAAGAGCCAGATGATGCTGGCAATTCATACTTAACAGAGATGGAAAAATTGGGTAATATT
GAGTTGAAACTTGAAAATATGAGAAAGATTTTAGAAACAGTCTGTCCTAAAAGAGGTACTGATTTT
TCCAATGATAAGAATCAAGAAGGAATCGCCAACCTTCTTAGTGAGGAGCTTCTTGTGTATGTACCA
GTGGTGGAGTTGAGGGAGAAGAGCTTACAGGTTGCTCCTGTCAGGTGA

SEQ ID NO: 188, sequência de polipeptídeo traduzida de PHDf_HD de Pinus radiata

MPKAGKGSSGSLRKKQALSNTASELRPKAAIQKTNNAGSISGRMRERKPRLHVQMIASILTKKRTG
EASESRNDSELEDKQYLSNSSKTNANFVAGKLVKKTLKKAPRLNLIKNDSQKKPSKSSASRSYAAQ
KPERPQQRKIRRTKWKGDSKNITKVKTVRTSNIDGLKRKQQRRKKKKKAQNVPQDEVSRAKRRIKY
LLLKMKLEQNLIDAYSGEGWKGQSREKIKPEKELQRAEKQILQCKLGIREAVHQLDLLSSEGSIAD
AVIDPEGRVFHEHIFCAKCKMTDALPDNDIILCDGACNRGFHQKCLDPPLATENIPPGDQGWLCKV
CECKLESLEAINAHLGTQFTVDNSWEGIFADAARIANGESTATGNGEEWPSDDSEDDDYDPEKQEI
RDETGYENNMSDSGKSSGWSESYNETDNESDASLTDNLNEYISKGKKRSRDTDEKNGQEAIIDLDT
DENDGSDMPVSGRRQRRDVDYKKLHDEMFGKDGPCEDEVSEDEDWGPNRRRRRAKQPDTNTTKDSS
NEEKEHVRKEADEKVAEHSQSSIKDKKHIVRLPANAVEKLRKVFAENELPSRSYKENLSQQLGLTF
RKVHMWFKNARYMSLKSRKEMPRNENCNETNLMGEGDMGPKKKIPKEKKAESTCLLSSVRAEYKSR
NTKRKHRRKNIKAPLSISSRSERKRTMCRLPANAVERFRQVFAENELPSMSCKMALSKQLDLPYRQ
VHMWFKNARYMSLKKKKMNRLKESAPVISSHTRRLEVESSKPNPQIKEPDDAGNSYLTEMEKLGNI
ELKLENMRKILETVCPKRGTDFSNDKNQEGIANLLSEELLVYVPVVELREKSLQVAPVR

SEQ ID NO: 189, sequência de ácido nucléico de PHDf_HD de Populus tremuloides

ATGGGTGATTCTGGAAAGACGTCAAAGCTGCAAGACTCGCACAAATGTTCACCTTCTGACACAGTG
ACTGGGTCGTTGCTGATTAAATCATTGAAGATAAAGAAGGACAGCAAAATATCTCCTAGAAAAGGC
CAAAAAACGAAAACAAAATCAAAACCAAAACCAATACCTCATCTAAAAACTATCATCAGTTCAGCT
GTATCAAAAAGAAAAGTTTCTCCTAAGGGCATTGGGAATGGCTCTACCAGTAGAAAACTGATTCAT
AGGAAAATTCTGCATAAAGCACTTGATAAGAAGGCCTCAAGGAAGGGGGCTTCCTCAGGGCTACAA
TTGTCAACTATTGATTCTAAGGGAAATGGAAAAAATGGTGATGAAGGTGCAATTAAAAAACTCAAG
AAGAGGAAGCCTAAGAAAAGGCAAAGGGACAAGGTAAAGCTAGATGAACCATCACGATTGCAGAGG
AGAGCAAGGTACCTGATGATTAAAATGAAGCTGGACCAGAACCTTATAGATGCTTACTCTGGAGAA
GGCTGGAAAGGTCAAAGTCGAGAGAAGATCAGGCCAGAAAAGGAACTACTGCGAGCAAGGAAGCAG
ATTTTGAAGTGTAAACTTGGTTTACGTGATATAATTCGGCAGGTGGATTCTCTTAGTACAGTAGGA
TGTATTGAAGAGACAGTTATGGCTCCAGATGGATCTGTTTCACATGAACATATATTCTGTGCGAAG
TGCAAATTGAATGAAGTTTCCCCAGATAATGATATTGTACTCTGTGATGGTACATGCAACTGTGCT
TTCCACCAAAAATGCCTTGAACCTCCCTTGGATACTGAAAGTATACCTCCAGGAGATCAGGGATGG
TTTTGCAAGTTCTGTGAATGCAGGATGGATATTATAGAAGCTATGAATGCTCATTTGGGAACCCAT
TTCTCAGAGGACAGCAGTTGGCAGGATATTTTCACAGAAGAAGCAGCAATTCCTGATGCTGGAAAT
GTGCTATTAAACCCCGAAGAAGAATGGCCTTCAGATGATTCTGAAGATGATAATTACGATCCAGAG
AGGAGAGACAACATCATGAGTGAAGCGGGTACTGATGATGATGCATCTGATGACATCAGCAGTTCA
ACCAGCTTGGGTTGGTCTTCAGATGGTGAGGTTTTTTTGGGATCCAGGAGGTGGGAGATGCATGGC
TTGGACTTCAGAAACAATTCTATTTATAGCAGTTTAGATTCTGATGAAACCAGTGATGGGGAAATT
GTTTGTGGCCGTAGGCAGAGAAGAGCAATTGACTATAAGAAGTTATATGATGAGGTGTTTGGAAAG
GATGCTCCTGCACACGAGCAAGCCAGTGAAGACGAAGATTGGGGTCCTGGCAAAAGAAAGCGACGG
GAGAAGGAGTCAAATGCAGCCAGCACCCTTATGACACTGTGTGAAAGTAAAAAAAAAAGTAAAAGT
GATGAGACCATTGAAGGTATGATGAACCTTCCCCCTCAAACTCGAAGGCCAATTTTCAGACTCCCC

FIGURA 19 CONT.

CCTGATGCAGTTGAGAAACTTCGCCAAGTATTTGTAGAAAATGAACTTCCATCTAGAACTGTCAAG
GAGAATCTGTCAAAAGAATTGGGCCTTGAACCTGGGAAGGTTAGCAAATGGTTCAAGAATTCCCGT
TATTTAGCTCTGAAGTCCCGGAAGGTAGAGAAAGGAGAACAACTTCATAACTCTAGTTCCAAAGTC
TCTGCTGAACCCACATTAAATGTCATGGAGGGCAAAACTGCTGATCTTTCACAGGATTCTTGGGAA
GAAACTGAGGTATGTATCCCAAAGAATTTGAAAAGGATCCTTCAGAGGAAGAAGCTGAAGTCAATA
AGCAGAAGTTTAAGGAAAAATGAACAGAAAAGAGGTTCCTACGAATCACCTACTAAAAGCAATGAG
ATGAATGCTGAGTGCAGTGATGATGTGAGCTTGAAGAAGTTGTTAAAAGCAAAACCAAAGGGAGTA
AAGAAAAAGGTTAATCCTATTTCAGTAGCGGCTGAATATGATATGGAGAGACTCTGTAGAGCGAAG
GCTATATTAGAGAAGATGAAGCAGAAAGTGGTGAAGCTACAAAATGGCAAAGCTAGGAAGTCATCC
AAAACCCGTCCACTTGAAGAATTTATCGTTTACATTCCCATTGCTGAGTTGAGGGGAAAAAATTGA

SEQ ID NO: 190, sequência de polipeptídeo traduzida de PHDf_HD de Populus tremuloides

MGDSGKTSKLQDSHKCSPSDTVTGSLLIKSLKIKKDSKISPRKGQKTKTKSKPKPIPHLKTIISSA
VSKRKVSPKGIGNGSTSRKLIHRKILHKALDKKASRKGASSGLQLSTIDSKGNGKNGDEGAIKKLK
KRKPKKRQRDKVKLDEPSRLQRRARYLMIKMKLDQNLIDAYSGEGWKGQSREKIRPEKELLRARKQ
ILKCKLGLRDIIRQVDSLSTVGCIEETVMAPDGSVSHEHIFCAKCKLNEVSPDNDIVLCDGTCNCA
FHQKCLEPPLDTESIPPGDQGWFCKFCECRMDIIEAMNAHLGTHFSEDSSWQDIFTEEAAIPDAGN
VLLNPEEEWPSDDSEDDNYDPERRDNIMSEAGTDDDASDDISSSTSLGWSSDGEVFLGSRRWEMHG
LDFRNNSIYSSLDSDETSDGEIVCGRRQRRAIDYKKLYDEVFGKDAPAHEQASEDEDWGPGKRKRR
EKESNAASTLMTLCESKKKSKSDETIEGMMNLPPQTRRPIFRLPPDAVEKLRQVFVENELPSRTVK
ENLSKELGLEPGKVSKWFKNSRYLALKSRKVEKGEQLHNSSSKVSAEPTLNVMEGKTADLSQDSWE
ETEVCIPKNLKRILQRKKLKSISRSLRKNEQKRGSYESPTKSNEMNAECSDDVSLKKLLKAKPKGV
KKKVNPISVAAEYDMERLCRAKAILEKMKQKVVKLQNGKARKSSKTRPLEEFIVYIPIAELRGKN

SEQ ID NO: 191, sequência de ácido nucléico de PHDf_HD de Saccharum officinarum contíguo de CA157855.1, CA261734.1, CA253314.1, CA220753.1, CA201958.1

ATGCATTCTTCGGAAAATAAACTGGTATGCTCCAACTCGGGAAGATCATCAAAAGGGAATGAGACC
TCTATGGAGTTGGTTCCTGTACCAAAAAGACCAACTAGACATGATGCTTCCCGCCAATGCAAATCT
GATTCTCCTTTGAAGCGATCACCAAGGAAAGCTAGAAATGCTACTCTGGCAAAAAGCATAAAGAAT
AAATATCATTGTAGTCCCCTCAAACAGCGGAGGGCTTCAGATTCTGTTTCTGGAAAAGTTGCGACA
GGACTAACTGTGAGGAGGAGGAAGAAAAGAAAAATGCAAAATACTGATGAGGCAACTCGCTTGGAA
CGAAGAGCGAGATATTTTCTAATCAAGATAAAATTGGAGCAGAATTTGCTAGATGCTTACTCTGGA
GATGGATGGAATGGGCAAAGCCGGGAAAAAATAAAGCCAGATAAGGATCTGCGGCGTGCCAGGAAA
CAAATCATAAAATGCAAAATTGCTATACGTGATATCATTCGCCAACTTGGTTTGTATACTTCTACT
GGGAGTCTTAATGACCCAGCAATGCCTCCAGATCAGTCTACCAATCCGGAACATACCATGTGCTCA
ACATGCAAGTCTCATGAATCATTTCCCAGCAATAAATTTATCTTCTGTGAAGGGCCCTGCAAAAGG
GCATATCATGAGAAATGTTTGGAACCTCCCTTAAACAAAAGCGTACTTCCAACAAGTAGCCATGGA
TGGCTTTGCAAATTCTGTTTGTGCAAAGTGAAAATTTTAGAAACTATTAATGCACATCTAGGAACA
AGTTTTACAGTGAAGTGCCCCTTCGAAGATATATTCAAGGAAGCAACGGAACAAATAGACTCCGAG
GATGCACTAGATGAAGATTGGCTTTCTGAGTATTCAGGTGATGAGGACTATGATCCTGATGAAAAT
GAGGACAGCGGCAACTGTATGGACAGCGGGGAGGAAATTATGTCTGATGATTCCGATGGTTCAGGA
AGCCCCCTTTATTCTCCAAATGACGATATTCCTGACTTCATATCAGCAGACTTAAATGACGTGGAA
GGGTTTTGTCACACCAATTTAGACTTAGGCATTGATGCCGTTGAAGATGATTTGGCACAGATCCTT
ACCTACCAAAGGCCAAGAAGAGATGTTGATTATAGAAGGCTTAATGAGGAAATGTTTGGAAAAATA
ATAGGAAATGAAGGACAGAGTGAGGATGAAGATTGGGGCCATGAAAGGAGAAAGAAAAGGACTTGT
TCAGGTGGTGCTGGGGATAATTCTGTGGGTTTCTCAAATGTTATATCTGAGGAAAAGAGCCAAAAG

FIGURA 19 CONT.

AAGGGGAGAAAACTTTTCAGGATTCCTCCTGCAGCTGTCGAGGTACTTCGCAAAGCTTTTGCTGAA
AATGAGCTTCCACCCCGGGATGTCAAAGAAAATCTCTCAAGAGAATTGGGAATTTCTTTTGAAAAG
ATTGATAAATGGTTCAAAAATACACGGTGTGCTGCTCTCAGAGATCGCAAGGCTGAAGGGAACAGT
CATAATACGGCTCCCAGCAAAAGCTCAAGAAATAAAGGAAAAGCTGGAATCTCAAGCCAGGCTGAA
AGGAATGGTCATGTTACTGGTCCCTGCAATAACTCGAGAACAAATGAAGAAAAATCGTTGACACTA
GCTCCTAATGCCCAGGGAGAGAAGAATGGGATTGCAGCAGATTATGTCTTACAAAAATCTGGTATA
TCGGGGAAGGTTGATTCAGTAGACAACTCTTGCTTGGTTCCCTTGTCTGAAAGCATCAATGTGCAT
ACACGATTGCAACAGAATCTTGAGATGAGGAAGATGGAATCAACTAGCAGTCCTGTGTGGCTGCAC
AATAAAGGAGGTTGCTTGTTTCCAACAGGCCAAGCCAAGGAGAGTACATTACCTACAGGCAAGCCT
TGTTTGCAGTCAGAAATAACCCATTCAACAATTAATGAAGTGAGCACTTTAGTGCAAGCTACTTCT
TGGATGGATGCTGGGTCGTGCGCTGAGGTGCAAGAAGCCACTCCTTGGGTGGACATCGGGGCCTCA
GATTACCAGCCTTTCCTGGACGTGATCGATGAGATGTGCGGACTTGAGTGCAGGCTGCAGAGACTG
AAGGAGAACATGCTCTCATCTGGCATAGACGGCAAAACCACAGGTGAGAGTGACATGGGAAACCAG
GCTGTGGTGCTCGTGCCGACCGCCGAGCTCAAGGAAAAAGCGCCGCATGGCAGTTTTTTTGGGCAT
TATTGCCCATAG

SEQ ID NO: 192, sequência de polipeptídeo traduzida de PHDf_HD de Saccharum officinarum

MHSSENKLVCSNSGRSSKGNETSMELVPVPKRPTRHDASRQCKSDSPLKRSPRKARNATLAKSIKN
KYHCSPLKQRRASDSVSGKVATGLTVRRRKKRKMQNTDEATRLERRARYFLIKIKLEQNLLDAYSG
DGWNGQSREKIKPDKDLRRARKQIIKCKIAIRDIIRQLGLYTSTGSLNDPAMPPDQSTNPEHTMCS
TCKSHESFPSNKFIFCEGPCKRAYHEKCLEPPLNKSVLPTSSHGWLCKFCLCKVKILETINAHLGT
SFTVKCPFEDIFKEATEQIDSEDALDEDWLSEYSGDEDYDPDENEDSGNCMDSGEEIMSDDSDGSG
SPLYSPNDDIPDFISADLNDVEGFCHTNLDLGIDAVEDDLAQILTYQRPRRDVDYRRLNEEMFGKI
IGNEGQSEDEDWGHERRKKRTCSGGAGDNSVGFSNVISEEKSQKKGRKLFRIPPAAVEVLRKAFAE
NELPPRDVKENLSRELGISFEKIDKWFKNTRCAALRDRKAEGNSHNTAPSKSSRNKGKAGISSQAE
RNGHVTGPCNNSRTNEEKSLTLAPNAQGEKNGIAADYVLQKSGISGKVDSVDNSCLVPLSESINVH
TRLQQNLEMRKMESTSSPVWLHNKGGCLFPTGQAKESTLPTGKPCLQSEITHSTINEVSTLVQATS
WMDAGSCAEVQEATPWVDIGASDYQPFLDVIDEMCGLECRLQRLKENMLSSGIDGKTTGESDMGNQ
AVVLVPTAELKEKAPHGSFFGHYCP

SEQ ID NO: 193, sequência de ácido nucléico de PHDf_HD de Vitis vinifera AM477372.2, AM488059.1

ATGCGTGGAACTGGAAAGAAAGCAGGACATCAGGAATCTGGAAAGTCTTGCTTTCCTAAAAGAAAT
ATCGGGCCTAAGCTGAATGCAGCATTGCAGATTAAAAATGGTAGTAAAATATCTCAGACCAGGAAA
TGCAAGCCAAAATCAAAATCTCATGCAAAGACAATTGGTGCAATCCTGTCTAAGAGAACAACTACT
GACTCTCCAAGCAAGGGAAGCAGGAGTGGATCCACAACCAGAAAACTGATTCACAAGAAAACCCTG
CATAAAGCCATTGATACTGAGTCTTCCAAGAAGGAGTCTTCATCAAAGCTTAAAGGTGAGAAGCCC
CCACAAATTAGCACAAACAAAAATGGGGAAACTGTGGACAAGAATGTTAAACCTCAAAAATTAAAG
AAAAGAGGAAAGCGGAAGCGGCGAAAGGATAATTCAGAGCTAGATGAAGCTTCTCGCTTGCAAAGA
AGAACCAGGTACCTCTTGATCAAAATGAAACTGGAGCAGAATCTTATTGATGCTTACTCTGGGGAA
GGTTGGAAAGGTCATAGTCGAGAAAAGATCAGGCCAGAGAAGGAACTACAAAGAGCTACGAAGCAA
ATATTGAAATGTAAACTTGGGATCCGTGATGCAATTCGCCAGCTGGAGTCACTGAGCTCCATAGGA
TGTATTGAAGATACTGCAATTGCTTCAGATGGATCTGTTTATCATGAACACATAATCTGTGCAAAG
TGCAAGTTGCGGGAGGCTTTCCCAGATAATGATATTATACTCTGTGATGGGACGTGCAATTGTGCT
TTCCACCAGAAATGTCTTGATCCTCCATTGGAAACTGAAAATATTCCTCCGGGAGACCAAGGTTGG
TTTTGCAAGTTCTGCGAGTGTAAGATGGAAATACTAGAAGCAATGAATGCACATCTTGGAACCCGC
TTCTCTGTGGACAGTACTTGGCAGGATATTTTCAAAGAAGAAGCTGCTTTGCCTGATGGTGGGAGT

FIGURA 19 CONT.

GCACTGCCCTATCCAGAGGAAGATTGGCCTTCTGATGATTCACAAGATCATGATTATGATCCTGAA
AGAAATGAAAATAGCTGCAGCATCAGCACTGCTGGCACTGAAGGAAATGCTTCAGATGATACAAAT
AGTTCTTTAAGCTTGAGTTGGTCGTTTGAAGACGAAATCCTTTCTGGATCCAAAAGATCAGGGATT
ATCAGTGCAGATTCTGATGAAACATCTGATTGTGAAATCATAAGTGGTCGTAGGCAACGAAGAGCT
GTTGATTATAGAAAGTTATATGATGAAATGTTTGGAAAGGATGCTCATGCAAATGAACAAGTTAGT
GAAGATGAAGACTGGGGTCCTGCTAACAAAAGGCGGAGAGAAAAGGAGTCTGATGCAGCCAGCACC
CTCATAACTCTATATGAAGGTGAGAAAAAGTTACCTAATGTGGAAACCATGGAAGCCAAACAAAAA
ATTTCTTCAGATCCACAAACTAAAAGGCCATTTTCCAGAATTCCACTTGATGCGGTTGAGAAGCTT
CGCCAAGCATTTGGGGAGAATGAACTTCCCTCTAGAGATGTGAGGGAGAATCTTGCAAAGCAGTTG
GGTCTTGATTATGAGAAGGTGAACAAATGGTTCAAAAATGCACGATATATAGCGCTAAAAACAAGA
AAGGCAGAAAGAGCAAAACAACTTCAGACTTCTCCCAGAATCTCCAAGGAATCCAGATCAGAAATC
GTGAAGGACAAAACTGTTGATCTAGTGGCATCAAGGGATAACTCATCAGCATCACTGGTCCGTGCA
CTAAAGAATTTGAAAAAGGTGCGGAGGAGAAAAAATCCAAAGCCAATAATGACCAGCCCTGTAAAG
AAAAAGCATCATAGAAGGGCTTTACTTGAGTCACCTACAAATGACAAGGTTACTATGGAGTTTGAC
GATGATGTGAGCTTGAAGAAACAACTGAAGCTTTTGAAAGAAAAAAGTAAGAGGGACAAGCAAAGG
GTCGACTTCAAGGAGGGAACTGGTGTACAAGATGCAGAGAAGGAGATGGAGAGACTCTGCCAAATT
AAGGATAAGATAGAGAAGTTGAAGCAGGTAATACTGAGACTTCAATGTGACAAAACTAATCAATGG
CAAGACCAGTCAGTGATCTATGTTCCTGTGGCAGAGCTAAGGGAGAAAGGTCGATTTTGCATCACA
TCT

SEQ ID NO: 194, sequência de polipeptídeo traduzida de PHDf_HD de Vitis vinifera

MRGTGKKAGHQESGKSCFPKRNIGPKLNAALQIKNGSKISQTRKCKPKSKSHAKTIGAILSKRTTT
DSPSKGSRSGSTTRKLIHKKTLHKAIDTESSKKESSSKLKGEKPPQISTNKNGETVDKNVKPQKLK
KRGKRKRRKDNSELDEASRLQRRTRYLLIKMKLEQNLIDAYSGEGWKGHSREKIRPEKELQRATKQ
ILKCKLGIRDAIRQLESLSSIGCIEDTAIASDGSVYHEHIICAKCKLREAFPDNDIILCDGTCNCA
FHQKCLDPPLETENIPPGDQGWFCKFCECKMEILEAMNAHLGTRFSVDSTWQDIFKEEAALPDGGS
ALPYPEEDWPSDDSQDHDYDPERNENSCSISTAGTEGNASDDTNSSLSLSWSFEDEILSGSKRSGI
ISADSDETSDCEIISGRRQRRAVDYRKLYDEMFGKDAHANEQVSEDEDWGPANKRRREKESDAAST
LITLYEGEKKLPNVETMEAKQKISSDPQTKRPFSRIPLDAVEKLRQAFGENELPSRDVRENLAKQL
GLDYEKVNKWFKNARYIALKTRKAERAKQLQTSPRISKESRSEIVKDKTVDLVASRDNSSASLVRA
LKNLKKVRRRKNPKPIMTSPVKKKHHRRALLESPTNDKVTMEFDDDVSLKKQLKLLKEKSKRDKQR
VDFKEGTGVQDAEKEMERLCQIKDKIEKLKQVILRLQCDKTNQWQDQSVIYVPVAELREKGRFCIT
S

SEQ ID NO: 195, sequência de ácido nucléico de PHDf_HD de Zea mays contíguo de EE162310, DN204182,CF057937

ATGCATTCTTCGGAAAATAAACTGGGTTGGAATGAGACCTCTATGGGGTTGGTTCCTGTACCAAAA
AGACCAGCTAGACCTGATGCTTCCCACCAATGCAAATCTGATTCCTTTATGAAGCGATCACCAAGG
AAAGTTAGAAATGCTACTCTGGCAAAAAGCATAAAGAGCAAATACCACTATAGTCCCCTCAAACAG
CGGAAGGGTTCAGATTCTGTTCCTGGGAAAATCGTAACAGGACTAACCGCAAGGAAAAAGAAGAAA
AGGAAAATCCAAATTACAGACGAGGCAACTCGTTTGGAACGAAGAGCGAGATATTTTCTAATCAAG
ATAAAACTGGAGCAGAATTTGCTAGATGCTTACTCTGGAGATGGATGGAATGGGCAAAGCCGAGAG
AAAATAAAGCCAGAGAAGGAACTGCAACGTGCCAGGAAACAAATCATAAAATGCAAAATTGCTATA
CGTGATATCATCCGCCAACTTGATTTGTATACTTCTACTGGGAGTGTTGATGACCCACTAATGCCT
ACAGATCAGTCCACCAATCCCGAACATACCATGTGCTCAACATGCAAGTCTCATGAATCATTTCCC
AGCAATAAAATTATCTTCTGCAAAGGGCCCTGCAAAAGGGCATGTCATGAGAAATGTTTGGAACCT
CCCTTAAACAAAAGCGTACTTCCAACAAGTAGCCATGGGTGGCTTTGCAAATTCTGTTTGTGCAAG

FIGURA 19 CONT.

GTGAGGATTTTAGAAACTATTAATGCACATCTAGGGACGAGTTTTACAGTGAAGTGCCACTTTGAA
GATATATTCAAGGAAACTACTGAACTAATAGACTCTGAGGATGCACTAGATGAAGATTGGCTTTCT
GAGTATTCAGGTGATGAGGACTATGATCCTGATGAAAATGAGGCCAGTGGCGACTGTATGGACAGC
GGGGAGAAGATTATGTCTGATGATTCCAATGGTTCAGGAAGCCCCCTTTATTCTCCAAATGACGAT
ATTCCTGACTTCATATCAGCAGACTTAAATGTTGTGGAAGGGTTTTGTCATACCAATTTAGATTTA
GGCATTGATGCCGTTGAAGATGATTTTGCACAGATCCTCACCTACCAAAGGCCAAGAAGAGATGTT
GATTATAGAAGGCTTAACGAGGAAATGTTTGGGAAAATAACCGGGAATGAAGAACAGAGCGAGGAC
GAAGATTGGGGCCATGAAAGGAGAAAGAAAAGGACGCATTCAGGTGTTGCTGGGGATAATTCTGTT
GGTTTCTTGAACGTTATATCTGATGAAAAGAGCCAGAAGAAGGGGAGAAAACTTTTCAGGATTCCT
CCTGCAGCTGTTGAGGTACTTCGCAGAGCTTTTGCTGAAAATGAGCTTCCACCCCGGGATGTTAAA
GAAAATCTCTCAAGAGAATTGGGAATTTCTTTTGAAAAGATTGATAAATGGTTCAAAAATACACGG
TGTGCTGCTCTCAGAGATCGGAAGGCTGAAGGAAACAGTCATAATACAGCTCCCAGCAAAAGCTCA
AAAAATAAAGGAAAAGCTGGAATCTCAGGCAAGACTGGAAGGAATGGTCATGTTACTGGTCCCTGC
AATAATTTGAGAACAAATGAAGAAAAAACAGGTATATCGGGGAAGTTTGATTCAGGAGACAACTCT
TGTTTGGTTCCCTTCTCTGAAGTCATCAATGTGCCTACGCGATTGCCACACAACTTTGAGATGAGG
AAGATGGAATCAACTAGAAGTCCTGCGAGGCTACACAATAAAGGAGGGTTCCTGTCTGCAACAGTC
CAAATTAAGGAGAGTACGTTGTTACCCACAGGCAAGCCTTGTTGGCAGTCAGAAATGAGCCATCCA
ACAACTAATGAAGTGAGCACTTCAGCGCAAGCTACTACTTGGATCGACGCGGGGGCGTGTGCCGAG
GAGCAAGAACCTACTCCTTGGGTGGACATCGGGGTCTCAGACTACCAGCCTTTCCTGGATGTGATC
GACGACATGTGCGGACTTGAGTGGAGGCTGCAGAGACTGAAGGAGAACATGCTCTCATCTAGCATA
GACGGCCAAACCGCCGCCAGTGAGAGTGACAAAAGAAACCAGACCGTGGTGCTAGTGCCAACCGCC
GAGCTCAAGGAAAAGCTGCCGCATGACGGTTTATTCGGGCATTATTGCCCATAG

SEQ ID NO: 196, sequência de polipeptídeo traduzida de PHDf_HD de Zea mays

MHSSENKLGWNETSMGLVPVPKRPARPDASHQCKSDSFMKRSPRKVRNATLAKSIKSKYHYSPLKQ
RKGSDSVPGKIVTGLTARKKKRKIQITDEATRLERRARYFLIKIKLEQNLLDAYSGDGWNGQSRE
KIKPEKELQRARKQIIKCKIAIRDIIRQLDLYTSTGSVDDPLMPTDQSTNPEHTMCSTCKSHESFP
SNKIIFCKGPCKRACHEKCLEPPLNKSVLPTSSHGWLCKFCLCKVRILETINAHLGTSFTVKCHFE
DIFKETTELIDSEDALDEDWLSEYSGDEDYDPDENEASGDCMDSGEKIMSDDSNGSGSPLYSPNDD
IPDFISADLNVVEGFCHTNLDLGIDAVEDDFAQILTYQRPRRDVDYRRLNEEMFGKITGNEEQSED
EDWGHERRKKRTHSGVAGDNSVGFLNVISDEKSQKKGRKLFRIPPAAVEVLRRAFAENELPPRDVK
ENLSRELGISFEKIDKWFKNTRCAALRDRKAEGNSHNTAPSKSSKNKGKAGISGKTGRNGHVTGPC
NNLRTNEEKTGISGKFDSGDNSCLVPFSEVINVPTRLPHNFEMRKMESTRSPARLHNKGGFLSATV
QIKESTLLPTGKPCWQSEMSHPTTNEVSTSAQATTWIDAGACAEEQEPTPWVDIGVSDYQPFLDVI
DDMCGLEWRLQRLKENMLSSSIDGQTAASESDKRNQTVVLVPTAELKEKLPHDGLFGHYCP

SEQ ID NO: sequência de ácido nucléico de Arabidopsis thaliana de PHDf_HD_HAT3.1 NM_112838 At3g19510

ATGTATAAAGCAGTAAGTAAGCGGGTAACTAGAAGTAGTGGATCTGGTTTGAAGCAGACAAATGTA
GATAATGGAGGAGAGATTAGTCCAACTGTTGACAGAGTATCTGAACAAGGCAAATCATCAGAGGCA
GGGAGTCACATGCCAACTGATGCAAACGGAAATGGACATTTGCATCATGAGATTATGGATCATGGA
AAGGGGAATGAAGAACAGAAACCAACCCCTCAAACCGTGAAGAAGGATTCTAATACCAATACCAAA
TTTTCAGGGTCTCATCGTGAGCTTGTCATCGGTCTTCCTTGCCGTGGACAATTTGAAATCCACAAC
CGGTCCAGGGCATCTACGAGTTCGAAGAGGTTAGGAGGAGGTGGAGAAAGGAATGTGTTGTTTGCA
AGTCACAAGAGAGCTCAGAGATCTAAGGAAGATGCTGGACCATCTTCGGTTGTTGCTAATTCCACA
CCTGTAGGTCGACCAAAGAAAAAGAATAAGACTATGAACAAAGGACAAGTTAGAGAAGATGATGAG
TACACAAGAATCAAAAAGAAACTTCGATACTTTCTAAACCGGATTAACTATGAGCAAAGCCTAATT
GACGCCTATTCGTTAGAAGGCTGGAAAGGTTCAAGCCTGGAAAAGATAAGGCCAGAGAAAGAGCTT

FIGURA 19 CONT.

GAACGAGCCACAAAGGAAATTTTGCGGCGCAAGCTTAAAATTAGAGATCTTTTTCAGCACCTTGAC
ACACTCTGTGCTGAAGGAAGCCTTCCAGAGTCTTTATTTGATACTGATGGAGAGATTTCTAGCGAG
GATATCTTCTGTGCAAAATGTGGTTCAAAGGACTTGTCTGTTGATAATGATATCATACTATGCGAT
GGGTTTTGCGATCGAGGCTTTCACCAATACTGTCTTGAACCACCTCTGCGGAAAGAAGATATTCCT
CCGGATGATGAGGGTTGGCTATGCCCCGGATGTGATTGCAAAGATGACAGCTTGGACTTGCTTAAT
GATTCTTTAGGGACAAAATTTTCGGTCAGCGACAGTTGGGAGAAAATATTTCCTGAGGCAGCAGCA
GCGCTGGTTGGTGGAGGTCAAAATCTTGACTGTGATCTTCCTTCAGATGATTCTGATGATGAAGAG
TATGATCCAGATTGTTTGAATGATAATGAAAATGATGAAGATGGCAGTGATGACAATGAAGAATCT
GAGAACGAAGATGGTAGTTCTGATGAAACTGAATTCACATCTGCATCTGACGAAATGATTGAGTCA
TTTAAAGAAGGGAAAGATATAATGAAAGATGTAATGGCCCTTCCATCTGATGATTCTGAGGATGAT
GACTATGATCCTGATGCCCCGACTTGTGATGATGATAAAGAAAGTTCAAATTCTGATTGTACATCA
GACACTGAAGATCTTGAAACTTCCTTCAAAGGAGATGAGACTAATCAGCAAGCTGAAGATACACCA
CTTGAGGATCCTGGCAGACAGACGTCTCAGCTCCAAGGTGATGCAATTTTAGAGTCAGATGTTGGG
CTTGATGATGGTCCTGCCGGTGTATCCAGAAGGAGGAATGTTGAAAGGCTGGATTACAAAAAGTTG
TATGATGAGGAATATGACAATGTTCCGACTTCGTCAAGTGATGATGACGATTGGGATAAAACTGCA
AGAATGGGAAAAGAAGATTCCGAGTCGGAAGATGAAGGGGACACTGTACCTCTGAAACAGTCTTCA
AATGCAGAAGATCATACTTCTAAGAAACTTATACGGAAGTCGAAAAGGGCAGATAAAAAGGATACC
TTAGAAATGCCACAAGAAGGTCCTGGAGAAAATGGTGGTTCTGGTGAAATTGAAAAAAGCAGCTCT
TCAGCATGTAAACAGACAGATCCCAAAACTCAGAGACTGTACATATCTTTTCAAGAGAATCAATAT
CCAGACAAAGCCACGAAGGAGAGCTTAGCCAAAGAGCTACAAATGACAGTTAAGCAAGTTAATAAC
TGGTTTAAGCATAGGCGTTGGTCGATAAACAGCAAACCTTTAGTTTCAGAGGAAAATGTCGAGAAG
TTAAAAACAGGCAAGGAAGGTGAATGTGAGACATCTGTTGCTGGAAGTAGCAAACAAACTATGGAA
ACCGAGTCAGTTGCAGAGAAACCAACCAATACTGGATCACGGAAAAGAAGAAGAAAGTAG

SEQ ID NO: 198, sequência de polipeptídeo traduzida de PHDf_HD_HAT3.1 de Arabidopsis thaliana

MYKAVSKRVTRSSGSGLKQTNVDNGGEISPTVDRVSEQGKSSEAGSHMPTDANGNGHLHHEIMDHG
KGNEEQKPTPQTVKKDSNTNTKFSGSHRELVIGLPCRGQFEIHNRSRASTSSKRLGGGGERNVLFA
SHKRAQRSKEDAGPSSVVANSTPVGRPKKKNKTMNKGQVREDDEYTRIKKKLRYFLNRINYEQSLI
DAYSLEGWKGSSLEKIRPEKELERATKEILRRKLKIRDLFQHLDTLCAEGSLPESLFDTDGEISSE
DIFCAKCGSKDLSVDNDIILCDGFCDRGFHQYCLEPPLRKEDIPPDDEGWLCPGCDCKDDSLDLLN
DSLGTKFSVSDSWEKIFPEAAAALVGGGQNLDCDLPSDDSDDEEYDPDCLNDNENDEDGSDDNEES
ENEDGSSDETEFTSASDEMIESFKEGKDIMKDVMALPSDDSEDDDYDPDAPTCDDDKESSNSDCTS
DTEDLETSFKGDETNQQAEDTPLEDPGRQTSQLQGDAILESDVGLDDGPAGVSRRRNVERLDYKKL
YDEEYDNVPTSSSDDDDWDKTARMGKEDSESEDEGDTVPLKQSSNAEDHTSKKLIRKSKRADKKDT
LEMPQEGPGENGGSGEIEKSSSSACKQTDPKTQRLYISFQENQYPDKATKESLAKELQMTVKQVNN
WFKHRRWSINSKPLVSEENVEKLKTGKEGECETSVAGSSKQTMETESVAEKPTNTGSRKRRRK

SEQ ID NO: 199, sequência de ácido nucléico de PHDf_HD de Lotus japonicus AP006117.1

ATGATTAGTTCAACTGGAGAATTGACAAGCCATAACAACAGTAGCATTGAGCACATGGAAACTGAG
CAACCTGAATTGTCTGAAAAAACACCTCAGATAGGCTCTGAATGTCTGGACGGTGAACAAAGGGTG
CTTGGTACTGTGTCAACCAGTTCTGTTAATAATGAAAATTCAAATCAAGTTTCTGCCAATGTGACC
GAGAATTCTGTTTTTCAACTACCAGCAACACCTCAACATTGCTTAGAAAAGAGTGGTCAGACTGTA
ACAAGTCCACCCCTGGAAGAAAGCACCTTGAAACAAGTTGCTACTAATGTATCTAATGGTAAATCA
GAAAACAAATGTCAGACATCCTCTCAAGATGTTCAAAACGAACTTGTGGAAGCAAGTGATGTAATA
ACTGGCTCTTTTGTTGCGGACCAAACACAATCCATTCCTGCTCAAGTGAATACGAGTTCTGTGAAT
GAACTATTGGATATTCCTTCCGGAGATGTTGCAGTAGAAGTTGGTTCTGATAGTTTAGAAAAAAAG

FIGURA 19 CONT.

TCTAAAGGGTGCTTGGTAGACAGCGAACAAAGGGTACTTGGTACTGTGTTAACCAATTCTATTAAT
GATGGAAATTCAAATCAAGTTTCTGCCAATGTGACCGAGAATTCTGTTATTCAACTGCCAGCACCA
CCTCAACATTGCTTTGAAAAGAGTGGACAGACTGTAAAAAGTTCATCTCTGGAAGAAAGTACCTTA
AAACAAGTCTCCACTAATTTATCTAACAGTAAATCAGAAAACATATGTCAGACATCCTCTCAAGAT
GTTCAAAACGAACTTGTGGAAACAAGTGATGCAGTAACTGGCTCTGTTGTTGCAGACAAAACACAA
TCCATTCCTGCTCAAGTGAATACAAGTTCTGTGAATGAACTATTGGATATTCCTTCCAGAGATGTT
GCAGAAGAAGTTGGTTCTGATAGTTCAGAAAGGAAGTCTAAAAGCACAACTCCTTCACAATTGAGA
CATATAGGTAAAAGTAACTCAAAGTTATCGAAAAAGAAGTATATCCTAAGGTCGTTAGGAAGCAGT
GACAGAGCTTTGCGGTCCAGGACAAAAGAAAAACCTAAAGCGCCTGAACCAAGTAGTAACTCAGTT
GACGTTAATAATGATGGAGTGAAGAAGGAAAAAAGGAGGAAGAAGAAAAAAACAGGAGGGGAGGAA
AAGAAGGATCAATTTTATAAAATCAAAGCTCACCTAAGATATTTATTAAACAGAGTAAGCTATGAG
CAAAACTTAATTGATGCTTACTCTGGTGACGGCTGGAAAGGATACAGTATGGACAAGTTAAAGCCT
GATAAGGAGATTCAACGGGCTAAATCTGAAATTCTTCGACGTAAATTGAAAATTAGGGATCTATTT
CAAAACTTGGATTCCTTGTGTGCTGAAGGAAAGTTTCCAGAATCTTTGTTTGATTCTGAGGGAGAG
ATTGACAGTGAGGATATATTCTGTTCAATTTGCCAGACCAAAGAGTTGAGCACTGATAATGATATA
ATTCTCTGTGATGGTGCTTGTGACCGTGGATTTCACCAGCACTGCTTGGATCCTCCATTGTTAACT
GAAGACATTCCACCTGGCGATGAGGGTTGGTTATGCCCCGGATGTGATTGCAAAGATGACTGCGTT
GACCTTCTTAACGACTCCTTAGGAACACGTCTCTCTCTTAGTGACACCTGGGAGAAGGTTTTTCCT
GAAGCAGCAGCTGTTACTGGAAATAATATGGATCAGGGACTTCCTTCTGATGATTCTGATGATGAT
GATTATAACCCTAATGGTAAAGAGGATGTGGAGGTTGAAGGAGGTGAATCAAGTTCTGATGAATCT
GAATATGCTTCTGCTTCTGAGAAATTGGAAGATTCACACCATGAGGATCCATACTTAGGCCTTCCT
TCTGAAGATTCAGAGGACAATGATTATGATCCTAGTGCTCCAGATCTTGACAATAAAGTCACAGAA
GAAAGTTCAAGCTCTGATTTCACATCTGACTCTGAGGATCTTGCTGCTGCTATTGAGGATAACACG
TCCCCTGGACATGTTAAACAGACTAAAAGCAGACAAAAAAGTAAAGTTGGTAAGAACCCATCAATG
GTTGATGAAGTGTCATCTTTACGAGAGCCAGAACTTGAGGAAGAGGGTTTCACTCCAGTTTCTGGG
AAAAGAAATGTGGAAAGGCTGGACTACAAAAAGCTGTATGATGAGACATACAAAAGTGATACTAGT
GAAGATGAAGAATGGACTGCCAGTGCCACTCCAAGTAGAAAAAAGAAACTCTGTGGCAAAATGACT
CCAGTGTCATCAGATGGAAAGGCCTCAAATAATTCTAGACATACTCCTGAAAGGAACACCCAACAA
GATAAAGTTGAAAATACAAATAATTCACCCACTAAATCACTTGAAGGCTGTCTGGAATCTGGCTCA
AGGGATAAAAAGCCTAGGTCTTCAACACGTCAAAGACTTGGAGACGTTGTAGTGCAGAGACTTCAC
AAATCTTTTAAAGACAATCAGTATCCGGATCGAGCCACAAAAGAAAGCTTGGCAGAAGAACTGGGG
CTCACTTTTTTCCAGGTTGACAAATGGTTTGGCAATGCTCGATGGGGCTTTCGGCGTTCATCGCGT
ATGGGAGCTAGTCCAGGCGAATATGCTTCGCCACAGGCCACTGGCAGCGGACCTGAAAATACTGGG
GAGAGGGAACGCGAATTAGCTTCCCAGGAAGTCGGTGAAGAAAAATTGAAAACCCCCAGTCCTAGA
AAAAGGAAGCATTTGTCTGAACCGCAGGCATCTGAAGCACCGCAGATTATTGTTCTTGGCTTAGCA
GCATCTCCTGGCTCACCACGGGCTCATGCAGGCAATAAAAAGAAGACAGTGAAAAGGAAATGA

SEQ ID NO: 200, sequência de polipeptídeo traduzida de PHDf_HD de Lotus japonicus

MISSTGELTSHNNSSIEHMETEQPELSEKTPQIGSECLDGEQRVLGTVSTSSVNNENSNQVSANVT
ENSVFQLPATPQHCLEKSGQTVTSPPLEESTLKQVATNVSNGKSENKCQTSSQDVQNELVEASDVI
TGSFVADQTQSIPAQVNTSSVNELLDIPSGDVAVEVGSDSLEKKSKGCLVDSEQRVLGTVLTNSIN
DGNSNQVSANVTENSVIQLPAPPQHCFEKSGQTVKSSSLEESTLKQVSTNLSNSKSENICQTSSQD
VQNELVETSDAVTGSVVADKTQSIPAQVNTSSVNELLDIPSRDVAEEVGSDSSERKSKSTTPSQLR
HIGKSNSKLSKKKYILRSLGSSDRALRSRTKEKPKAPEPSSNSVDVNNDGVKKEKRRKKKKTGGEE
KKDQFYKIKAHLRYLLNRVSYEQNLIDAYSGDGWKGYSMDKLKPDKEIQRAKSEILRRKLKIRDLF
QNLDSLCAEGKFPESLFDSEGEIDSEDIFCSICQTKELSTDNDIILCDGACDRGFHQHCLDPPLLT
EDIPPGDEGWLCPGCDCKDDCVDLLNDSLGTRLSLSDTWEKVFPEAAAVTGNNMDQGLPSDDSDDD

FIGURA 19 CONT.

DYNPNGKEDVEVEGGESSSDESEYASASEKLEDSHHEDPYLGLPSEDSEDNDYDPSAPDLDNKVTE
ESSSSDFTSDSEDLAAAIEDNTSPGHVKQTKSRQKSKVGKNPSMVDEVSSLREPELEEEGFTPVSG
KRNVERLDYKKLYDETYKSDTSEDEEWTASATPSRKKKLCGKMTPVSSDGKASNNSRHTPERNTQQ
DKVENTNNSPTKSLEGCLESGSRDKKPRSSTRQRLGDVVVQRLHKSFKDNQYPDRATKESLAEELG
LTFFQVDKWFGNARWGFRRSSRMGASPGEYASPQATGSGPENTGERERELASQEVGEEKLKTPSPR
KRKHLSEPQASEAPQIIVLGLAASPGSPRAHAGNKKKTVKRK

SEQ ID NO: 201, sequência de ácido nucléico de PHDf_HD_HAZ1 de Oryza sativa AB081340

ATGGATAAGACAACTACTTCTGATCTTGTTTTGGACAACGACAACATTGGGAGTAATGCTGGTTCT
GCACAGGAGCCTCTTACTACCAATGGTAAGACTAGTGGGGTCAGAAATAGATACAAACAAACTGTT
AAAAGGGGAAGGAAAGGCTCCCAAATATCACCAAGTAAAACATATCCCCTGAGATCTTCACATAGT
AATGTCAGGGTGCTTCGCTCTGCCTCAAAAAAGAAGAATGAGACACCTATTGTGCCCACAAATGAT
AATACTGCTGTTCAACGAGTTGCGAAGAAAAGGAAAAGAAGCAAACCCTTGAGACCTGCACCTAGC
AGGGTGCTTCGCTCTACCTCGGAAAAGAAGAATAAGGCACATAATGAGCTTCTAAATGATGGTGCT
GGTGTTCAACCAGCCGAAAAGAAAAGAAAGGTAGGCAGACCCCCAAAAGGAGGAACTCCCAAGGAC
GATTACCTCATGATCCGAAAGCGAGTTAGATATGTTTTGAATCGAATGAACTATGAACAAAGTTTA
ATTCAAGCCTATGCCAGTGAAGGTTGGAAAGGTCAAAGCTTGGAAAAAATAAGACCTGAGAAGGAG
CTTGAACGAGCCAAGGTGGAAATTCTGCGATGCAAGTCAAGAATACGGGAAGCTTTTCGGAATTTG
GATTCCCTCCTATCTGAGGGAAAGCTGGATGAATCTATGTTTGATTCTGCTGGAGAAATATCTAGT
GAAGATATTTTCTGTGCTGCATGTGGTTCAAAGGATGTTACATTAAAAAATGACATAATTCTTTGT
GATGGAATCTGCGACAGAGGGTTCCACCAGTATTGTTTAAACCCTCCTTTGCTTGCTGAAGATATT
CCACAGGGGGATGAAGGATGGCTTTGCCCTGCATGTGATTGCAAAATTGACTGCATAGATGTATTA
AATGAACTCCAAGGGGTCAAACTCTCTATCCATGACTCTTGGGAGAAAGTTTTTCCTGAGGCAGCA
TCCTTTTTGAATGGTTCTAAGCAAATTGATGCGTCCGATCTTCCATCAGATGATTCAGCAGACAAT
GATTATGACCCTACTTTGGCTCAAGGGCACAAGGTGGATGAAGAAAAATCTTCTGGAGAAGATGGT
GGTGAAGGATTAGATTCTGATGACTCATCTTCTGAGGATTCTGAATCTTCTGAGAAGGAAAAGTCT
AAAACTTCACAGAATGGAAGGACAGTTGATGATCTTGGTTTGCCTTCTGAGGATTCAGAGGATGGT
GACTTTGATCCAGCAGGTCCAGATTCCGACAAAGAACAAAATGATGAATCAAACTCAGATCAGTCA
GATGAATCAGATTTTACATCTGACTCTGATGATTTTTGTGCTGAGATTGCTAAATCCTGTGGTCAG
GATGAAATCTCAGGCCCTTCATCATCACAAATCAGAACAGTCGATCGCACCGATGGAAGTGGTTTT
GATGGTGAGCCTAATGCAGAAAATTCGAATCTTGCATTTATGGAGACAGAGCTAGAGCAAGACATG
GTGCTACCAATTTCTAGTAAACGACAAGTTGAACGTTTGGATTACAAAAAACTGTATAATGAGGCT
TATGGTAAAGCATCATCTGATTCAAGTGACGATGAAGAGTGGTATGGAAATAGTACACCCGAAAAA
GGTAATTTAGAAGACAGTGAAACAGATTCGCTTGCTGAATCACCTCAGGGTGGAAAAGGATTCTCT
AGAAGAGCACCAGTTAGGTACCATAATAATGAACATACTCCACAGAATGTAAGGCCCGGTGGTTCA
GTAAGCGACCAGCAAACTGAAGTACTTTGCTCCAATAGCAATGGTAGCACAGCCAAAAACAGACAT
TTTGGTCCTGCAATTAATCAGAAACTGAAGGCACATTTCAAGGAAGATCCATACCCTAGTCGTGCA
ACAAAAGAAAATCTGGCACAAGAGTTAGGCCTCACATTTAATCAGGTCACTAAATGGTTTTCCAGT
ACTCGTCATTACGCAAGAGTGGCAGCCACAAAAAAGGAGAACAACATTGAAAATCATACGGCTGAG
AATAATAACAATACCAACACTGTTGATAGCATACAACTGAGAGGATCCAATGATATTGTTAGTGTA
GATAGAAATGATATGGTCTCCGAGGAAAGGACAGGACAAAGTAATCTCAATGAAGGTACTCCATTA
AGATCAGATACCAGCTGTGGGCAGAGTGTGGCTGTGACTCCTATGGTACACCCAGAAAACCAGGGC
AATGATTCCTCAAGTAATGTTAGAACACCAAATGCTAAGTCTGCGGAGAAATTGATCCCTGGACTG
GAAAATTCAGATGAGGCTAGGCGAAAGGCGGTACAACGGGAACTGAGAAAGATGAAGACTGGCAGG
TGA

FIGURA 19 CONT.

SEQ ID NO: 202, sequência de polipeptídeo traduzida de PHDf_HD_HAZ1 de Oryza sativa

MDKTTTSDLVLDNDNIGSNAGSAQEPLTTNGKTSGVRNRYKQTVKRGRKGSQISPSKTYPLRSSHS
NVRVLRSASKKKNETPIVPTNDNTAVQRVAKKRKRSKPLRPAPSRVLRSTSEKKNKAHNELLNDGA
GVQPAEKKRKVGRPPKGGTPKDDYLMIRKRVRYVLNRMNYEQSLIQAYASEGWKGQSLEKIRPEKE
LERAKVEILRCKSRIREAFRNLDSLLSEGKLDESMFDSAGEISSEDIFCAACGSKDVTLKNDIILC
DGICDRGFHQYCLNPPLLAEDIPQGDEGWLCPACDCKIDCIDVLNELQGVKLSIHDSWEKVFPEAA
SFLNGSKQIDASDLPSDDSADNDYDPTLAQGHKVDEEKSSGEDGGEGLDSDDSSSEDSESSEKEKS
KTSQNGRTVDDLGLPSEDSEDGDFDPAGPDSDKEQNDESNSDQSDESDFTSDSDDFCAEIAKSCGQ
DEISGPSSSQIRTVDRTDGSGFDGEPNAENSNLAFMETELEQDMVLPISSKRQVERLDYKKLYNEA
YGKASSDSSDDEEWYGNSTPEKGNLEDSETDSLAESPQGGKGFSRRAPVRYHNNEHTPQNVRPGGS
VSDQQTEVLCSNSNGSTAKNRHFGPAINQKLKAHFKEDPYPSRATKENLAQELGLTFNQVTKWFSS
TRHYARVAATKKENNIENHTAENNNNTNTVDSIQLRGSNDIVSVDRNDMVSEERTGQSNLNEGTPL
RSDTSCGQSVAVTPMVHPENQGNDSSSNVRTPNAKSAEKLIPGLENSDEARRKAVQRELRKMKTGR

SEQ ID NO: 203, sequência de ácido nucléico de PHDf_HD_PRHP de Petroselinum crispum L21975

ATGGAAGAGATCTCAGACCCTAAACCAAATGCATTAGAGCAAGTACTTCCTACTGTCCCAAATGGT
AAATGTACTGCACCGGTGCAAATGGAAAGTTTGGCTGTTGATGTTCAAAAAGTATCAGGTGAAGCA
AAGGTGAGGATATGTTCCTGTTGGTGCGAGATAGTACGTTCTCCAGAAGATTTAACTAAACTTGTG
CCTTGTAATGATTTTGCAGAAGATATAAAATTGTTTGATTCTGATCCCATGCAACAAGAAGCTGAA
AGCTCCATTGGAATACCATTGATCCCTAAACAAGTTACAATGTCCCATAACCATGACCATGAATCT
GGTTCTGAAATGGTCAGTAACGAAGTAATGCAAGAAAACCATGTAATTGCCACAGAAAATACTTAC
CAGAAATCTGATTTTGACCGTATAAATATGGGACAAAAAGAAACCATGCCAGAAGAAGTGATCCAT
AAGTCTTTCTTGGAATCCTCAACTTCCTCGATAGATATACTATTAAATAATCATAACAGTTATCAA
TCAGGACTGCCACCAGAGAATGCAGTAACAGATTGCAAGCAAGTCCAATTGGGACACCGAAGTGAT
GATGCTATAAAGAATTCTGGTCTTGTGGAATTGGTAATAGGTCAGAAAAATGTTGCCAAGAGTCCT
AGCCAGTTGGTAGAAACGGGGAAAAGAGGTCGTGGTCGGCCCAGGAAAGTACAAACTGGCCTTGAG
CAATTGGTAATAGGTCAGAAAACTGCTGCCAAGAGCTCTAGCCAGTTAGGTGACACAGGGAAAAGA
TCTCGTGGTCGGCCCAGAAAAGTACAAAATTCTCCAACTTCTTTCTTGGAGAATATAAATATGGAA
CAAAAAGAAACAATACCAGAACAAGTGACCCAAAATTCTATCTTGGAGTCCTTGACTATCCCGACA
GACAATCAGTCAAGAACTTATAACAGTGATCAATCAGAACTGCCACCAGAGAATGCAGCAAAAAAT
TGCAATCACGCCCAATTTGGACATCAAAGTGATGATACTACAAAGATTTCTGGATTCAAGGAATTG
GTAATAGGTCAGGAAACAGTTGCAAAGAGTCCTAGCCAGTTGGTAGACGCAGGAAAAAGAGGTCGT
GGTCGGCCCAGAAAAGTACAAACTGGTCTTGAGCAATTGGTACCAGTTCAGGAAACTGCGGCCAAG
AGCTCTAGCCAGTTGGGAGACACAGGGAAAAGATCTCGCGGCCGGCCCAGGAAAGTACAAGATTCT
CCAACTTCTTTAGGGGGTAATGTTAAAGTGGTACCAGAGAAAGGAAAAGATTCTCAGGAATTGTCG
GTAAACAGCAGCAGATCATTGCGCTCAAGGTCACAAGAGAAATCTATCGAACCCGATGTCAATAAT
ATTGTGGCAGATGAAGGTGCTGACAGAGAAAAACCAAGGAAGAAGAGGAAAAAGCGAATGGAGGAA
AATAGGGTTGATGAATTCTGTAGAATAAGGACACACCTAAGGTACTTACTACACAGAATAAAATAC
GAGAAGAATTTCCTTGATGCTTACTCTGGGGAGGGATGGAAGGGACAAAGCCTAGACAAAATAAAA
CCTGAGAAGGAGCTTAAACGAGCAAAAGCTGAAATTTTTGGACGCAAATTGAAAATTAGAGATCTA
TTTCAGCGCTTAGATTTAGCACGTTCCGAGGGAAGGCTTCCAGAAATTCTATTTGATTCTCGAGGA
GAAATTGACAGCGAGGATATATTCTGCGCGAAATGTGGATCCAAGGATGTGACACTCAGTAATGAT
ATCATCCTTTGCGATGGTGCATGTGATCGTGGATTTCACCAATTTTGTTTGGACCCACCATTGCTG
AAAGAATATATTCCTCCTGACGACGAGGGTTGGCTTTGCCCGGGATGCGAGTGCAAAATTGACTGC
ATCAAGTTACTTAATGATTCCCAAGAGACAAATATTTTACTCGGTGACAGCTGGGAGAAAGTTTTT
GCTGAGGAGGCAGCAGCAGCAGCTTCCGGAAAAAATCTAGATGATAATTCTGGACTGCCTTCAGAT

FIGURA 19 CONT.

GATTCCGAGGATGATGATTATGATCCTGGGGGCCCTGATTTGGATGAGAAGGTGCAGGGAGATGAT
TCGAGTACGGATGAATCTGATTATCAGTCGGAATCAGATGATATGCAAGTTATACGTCAAAAAAAT
TCGCGCGGGCTTCCTTCAGATGATTCCGAAGATGATGAATACGATCCTAGTGGCTTAGTTACTGAT
CAGATGTATAAGGACAGCTCATGTTCTGATTTTACATCTGATTCTGAGGATTTTACAGGTGTGTTT
GACGATTATAAGGATACTGGTAAAGCTCAAGGGCCACTGGCATCGACCCCAGATCACGTCAGGAAC
AATGAGGAAGGGTGTGGGCATCCTGAGCAAGGTGATACTGCCCCCCTGTATCCGAGAAGACAAGTA
GAGAGTTTGGACTACAAAAAGCTAAATGATATTGAGTTCTCTAAAATGTGTGACATCCTTGATATT
CTTTCAAGTCAACTTGATGTAATTATTTGTACTGGTAACCAGGAAGAGTATGGGAACACATCTTCT
GATTCAAGTGATGAAGATTATATGGTTACTTCTTCACCAGATAAAAATAATTCTGATAAAGAAGCT
ACAGCAATGGAGCGTGGAAGAGAGAGCGGTGATTTGGAGCTAGACCAAAAAGCGAGAGAATCCACT
CATAATAGGAGATATATAAAGAAGTTTGCTGTTGAAGGCACAGATTCTTTTCTTTCTAGATCGTGT
GAAGACTCTGCAGCACCTGTTGCTGGTAGTAAAAGTACTTCAAAAACATTGCACGGGGAACATGCA
ACTCAGAGACTCCTCCAATCCTTCAAGGAAAATCAATACCCTCAACGAGCTGTGAAAGAGAGTTTG
GCTGCAGAATTAGCACTTTCAGTCCGACAGGTTAGCAACTGGTTTAATAATAGACGTTGGAGCTTT
CGCCACTCATCACGTATTGGATCGGATGTAGCTAAATTTGATTCAAACGATACTCCCCGTCAGAAA
AGCATTGACATGTCTGGACCCAGCTTGAAATCGGTCTTGGACAGTGCTACTTACAGTGAAATTGAA
AAGAAGGAACAAGATACGGCAAGCCTTGGTTTGACAGAAGGTTGTGACAGATACATGACGCTGAAT
ATGGTTGCAGATGAAGGCAATGTGCACACTCCTTGTATTGCAGAAACTAGGGAAGAAAAAACCGAA
GTCGGTATAAAGCCACAACAGAACCCACTATAA

SEQ ID NO: 204, sequência de polipeptídeo traduzida de PHDf_HD_PRHP de Petroselinum crispum

MEEISDPKPNALEQVLPTVPNGKCTAPVQMESLAVDVQKVSGEAKVRICSCWCEIVRSPEDLTKLV
PCNDFAEDIKLFDSDPMQQEAESSIGIPLIPKQVTMSHNHDHESGSEMVSNEVMQENHVIATENTY
QKSDFDRINMGQKETMPEEVIHKSFLESSTSSIDILLNNHNSYQSGLPPENAVTDCKQVQLGHRSD
DAIKNSGLVELVIGQKNVAKSPSQLVETGKRGRGRPRKVQTGLEQLVIGQKTAAKSSSQLGDTGKR
SRGRPRKVQNSPTSFLENINMEQKETIPEQVTQNSILESLTIPTDNQSRTYNSDQSELPPENAAKN
CNHAQFGHQSDDTTKISGFKELVIGQETVAKSPSQLVDAGKRGRGRPRKVQTGLEQLVPVQETAAK
SSSQLGDTGKRSRGRPRKVQDSPTSLGGNVKVVPEKGKDSQELSVNSSRSLRSRSQEKSIEPDVNN
IVADEGADREKPRKKRKKRMEENRVDEFCRIRTHLRYLLHRIKYEKNFLDAYSGEGWKGQSLDKIK
PEKELKRAKAEIFGRKLKIRDLFQRLDLARSEGRLPEILFDSRGEIDSEDIFCAKCGSKDVTLSND
IILCDGACDRGFHQFCLDPPLLKEYIPPDDEGWLCPGCECKIDCIKLLNDSQETNILLGDSWEKVF
AEEAAAAAASGKNLDDNSGLPSDDSEDDDYDPGGPDLDEKVQGDDSSTDESDYQSESDDMQVIRQKN
SRGLPSDDSEDDEYDPSGLVTDQMYKDSSCSDFTSDSEDFTGVFDDYKDTGKAQGPLASTPDHVRN
NEEGCGHPEQGDTAPLYPRRQVESLDYKKLNDIEFSKMCDILDILSSQLDVIICTGNQEEYGNTSS
DSSDEDYMVTSSPDKNNSDKEATAMERGRESGDLELDQKARESTHNRRYIKKFAVEGTDSFLSRSC
EDSAAPVAGSKSTSKTLHGEHATQRLLQSFKENQYPQRAVKESLAAELALSVRQVSNWFNNRRWSF
RHSSRIGSDVAKFDSNDTPRQKSIDMSGPSLKSVLDSATYSEIEKKEQDTASLGLTEGCDRYMTLN
MVADEGNVHTPCIAETREEKTEVGIKPQQNPL

SEQ ID NO: 205, sequência de ácido nucléico de PHDf_HD_HOX1a de Zea mays X67561

ATGGAAAAAAATATAGCTCACTGTCCTGTTGAGGGCAATGGTGAGATTGAAAATGGTGCGAGCTCT
AGTCAGAATCCTGAGTCTCTTGAGCATTCAGTTTTGCTGTCTACATCTCAAACGATGCCAAATAAC
TTGGGGATAAGGAAAAACTATAAAAGGGCAGCAAACAGAGGTAAAAAAGGTTCCCAAGGACTAACA
GGCCAAGCATATACTTTGATGTCATCTAATAGTGATGTCAGGGTGCTTCGCTCCACATCAAGTTCA
AAGACAACATCTACTGAGCATGTTCAGGCTCCAGTACAACCGGCTGCCAAGAGAAGAAAAATGAGC
AGAGCCTCAAACAAAAGTTCGACTGATGAGTTTTCACAAATTCGTAAACGAGTTAGATATATTTTG

FIGURA 19 CONT.

AACCGAATGAATTATGAGCAAAGCCTAATTGAAGCTTATGCTAGCGAAGGCTGGAAAAATCAAAGT
TTGGATAAGATAAGACCTGAGAAGGAGCTTGAACGAGCCAAATCAGAAATCTTACGATGCAAATTG
AGAATACGAGAAGTCTTCCGGAATATTGATTCTCTTCTCTCCAAGGGGAAAATTGACGAAACTTTA
TTTGATTCTGAAGGAGAAATATCTTGCGAAGATATCTTTTGTTCCACTTGTGGTTCGAATGACGCT
ACATTGGGTAATGATATCATTCTCTGCGATGGAGCTTGTGACAGAGGGTTCCACCAGAACTGTTTA
AATCCTCCCCTGCGTACTGAAGATATCCCCATGGGAGATGAAGGATGGCTCTGCCCAGCATGTGAT
TGCAAGATAGATTGTATAGATCTAATAAATGAACTCCATGGGAGTAACATCTCCATTGAGGACTCT
TGGGAGAAAGTTTTTCCAGATGCAGCTGCTATGGCAAATGACTCTAAACAAGATGACGCATTTGAT
CTTCCATCAGATGACTCAGACGACAATGACTTTGACCCCAATATGCCTGAAGAGCATGTGGTTGGT
AAGGATGAAGAATCGTCTGAAGAGGATGAGGATGGAGGATCAGACTCTGATGACTCAGACTTTTTG
ACATGTTCTGATGATTCGGAACCTTTGATAGACAAGAAGGTTGATGACCTCAGATTACCTTCTGAA
GATTCAGAGGATGATGACTATGATCCAGCAGGTCCTGATTCAGATAAAGATGTCGAAAAGAAGTCA
AGTTCTGATGAATCTGACTTCACATCTGACTCAGATGATTTTTGTAAGGAGATTTCAAAATCTGGA
CATGATGAAGTTTCATCACCTCTATTACCCGATGCCAAGGTGGGTGATATGGAAAAGATTACTGCT
CAAGCTAAAACAACAAGTTCTGCTGACGACCCTATGGAGACTGAAATAGACCAGGGTGTGGTTTTA
CCAGATTCAAGGAGACGGCAAGCTGAACGGTTGGACTACAAAAAGTTGTATGATGAGGCATATGGT
GAAGCATCATCTGATTCTAGTGATGATGAAGAATGGTCTGGGAAGAACACGCCAATAATAAAAAGC
AATGAAGAGGGTGAAGCCAATTCTCCTGCGGGAAAAGGCTCCAGAGTTGTGCACCATAATGATGAA
CTGACTACACAAAGCACTAAAAAAAGCTTGCATTCTATTCATGGTTCAGTAGATGAGAAACCTGGA
GACCTTACTTCTAATGGCAGCAACAGCACAGCCAGGAAAGGACATTTTGGCCCAGTAATTAATCAG
AAGCTACATGAACATTTTAAGACACAACCATACCCTAGTCGTTCTGTTAAAGAAAGTCTGGCAGAA
GAACTAGGGTTAACATTTCGTCAGGTTAACAAATGGTTTGAAACTAGGCGTCATTCCGCAAGGGTA
GCTTCTTCCAGGAAAGGCATCAGTCTAGATAAGCATAGCCCCCAAAATACTAATAGTCAAGTGACA
GCTAGTATGGAACCTAAAGAACCTGAGGGGACTGTGGTGGAAGAATCTAATGTATGCCTAAATGGG
GGTACTACCATCTCTAAAGAAGCAGTGTCTTCAAAAGTTGGATCAAGAACCCCTGGAAGTGATGTG
GGAGGGTCAAAAGTTGATTCTGCTGAGGATCAGAATCCGGGGCCTGATCTTGCAGAGAAGGCAAGG
CAAAAGGCAATACAGCAGGAGTTGAGGAAGAAAAAAATGGGACGATGA

SEQ ID NO: 206, sequência de polipeptídeo traduzida de PHDf_HD_HOX1a de Zea mays

MEKNIAHCPVEGNGEIENGASSSQNPESLEHSVLLSTSQTMPNNLGIRKNYKRAANRGKKGSQGLT
GQAYTLMSSNSDVRVLRSTSSSKTTSTEHVQAPVQPAAKRRKMSRASNKSSTDEFSQIRKRVRYIL
NRMNYEQSLIEAYASEGWKNQSLDKIRPEKELERAKSEILRCKLRIREVFRNIDSLLSKGKIDETL
FDSEGEISCEDIFCSTCGSNDATLGNDIILCDGACDRGFHQNCLNPPLRTEDIPMGDEGWLCPACD
CKIDCIDLINELHGSNISIEDSWEKVFPDAAAMANDSKQDDAFDLPSDDSDDNDFDPNMPEEHVVG
KDEESSEEDEDGGSDSDDSDFLTCSDDSEPLIDKKVDDLRLPSEDSEDDDYDPAGPDSDKDVEKKS
SSDESDFTSDSDDFCKEISKSGHDEVSSPLLPDAKVGDMEKITAQAKTTSSADDPMETEIDQGVVL
PDSRRRQAERLDYKKLYDEAYGEASSDSSDDEEWSGKNTPIIKSNEEGEANSPAGKGSRVVHHNDE
LTTQSTKKSLHSIHGSVDEKPGDLTSNGSNSTARKGHFGPVINQKLHEHFKTQPYPSRSVKESLAE
ELGLTFRQVNKWFETRRHSARVASSRKGISLDKHSPQNTNSQVTASMEPKEPEGTVVEESNVCLNG
GTTISKEAVSSKVGSRTPGSDVGGSKVDSAEDQNPGPDLAEKARQKAIQQELRKKKMGR

SEQ ID NO: 207, sequência de ácido nucléico de PHDf_HD_HOX1b de Zea mays X92428

ATGGAAAAAAAGACAGTTCACGTTCTGGCTGAGGGTAATGGTGAGATTGAAAATGGTACGAGCTCT
AGTCAGAATCCTGAGACCCTTGAGCATCCAGTTTTGCTGTCTGCATCTCAAACTGTGCCAAATAAC
TTGGGGATAAGGAAAAACTATAAAAGGGCAGCAAACAGAGGTAAAAAGGGTTCCCAAGGACTAACA
GGTCAGGCATATACATTGAGGTCATCTGATATTGATGTCAGGGTGCTTCGCTCTACATCAGGTTCA

FIGURA 19 CONT.

AAGACAACATCTACTGAGCATGTGCAGCCTCCAGGACAACCAGCTGCCAAGAGAAGAAAAAGGAGC
AGAGCCTCAAACAAAAATTCGACAGATGAGTTCTCACAAATTCGTAAACGGGTTAGATACATTTTG
AACCGAATGAATTATGAGCAAAGCCTAATTGAAGCTTATGCTAGTGAAGGCTGGAAAAATCAAAGT
TTGGATAAGATAAGACCTGAGAAGGAGCTTGAACGAGCCAAATCAGAAATCTTACGATGCAAATTG
AGAATACGGGAAGTCTTCCAAAATATTGATTCTCTTCTGTCCAAAGGGAAAATTGACGAATCTTTA
TTTGATTCTGAAGGAGAAATATCTTGTGAAGATATCTTTTGTGCTACTTGTGGTTCGAAAAATGCT
ACATTGGGTAATGATATCATTCTCTGTGATGGAGCTTGTGACAGAGGGTTCCACCAGAACTGTCTA
AATCCTCCTCTGCGTACTGAAGATATTCCCATGGGAGATGAAGGATGGCTCTGCCCAGCATGTGAT
TGCAAGTTAGACTGTATAGATCTAATAAATGAACTCCAGGGCAGTGACATCTCCATTGAGGACCCT
TGGGAGAAAGTTTTTCCAGAGGCAGCTGCTATGACCAATGGCTCTAAGCAAGATGAACGATTTGAT
CTTCCATCTGATGATTCGGATGACAATGATTTTGACCCCAATATGCCTGAAGAGCATGTGGCTAGT
AAGGAAGAAGGATCTTCTGAAGAGGAAGAGGATGATGACGGAGGATCAGACTCTGACGACTCTGAC
TTCTTAACCTGTTCTGATGATTTGGAACCTTTGATAGATAAAAAGAAGGTTGATGACCTGGGATTA
TCTTCTGAAGATTCAGAGGATGATGACTATGATCCAGCAGGTCCTGATTCGGATAAAGATGTCGAA
AAGAAGTCAAATTCTGATGAGTCTGATTTCACGTCTGACTCGGATGATTTTTGTAAGGAGATTAAA
AAATCTGGTGGACATGATGAAGTTTCATCGCCTCCATTACCCGATGTCAAGGTGGGTGATATGGAA
AAGAACACTGCTCAGTCTAACACAACAAGTTCTGCTGACGACCCTATGGAGACTGAAATAGACCAG
AGTGTGGTTTTACCAGTATCAAGGAGACGTCAAGCTGAACGGTTGGACTACAAAAGGCTGTATGAT
GAGGCATATGGTGAAGCATCATCTGATTCTAGTGATGAAGAAGAATGGTCTGGGAAGAACACACCA
ATAAAAAGCAATGAAGAGGGTGAAGTTGGTTCTCCTGCAGGAAAAGGCTCTAGAGTTGCGCACCAT
AATGAACTGACTACACAAAACACTAAAGAAAGCTTGCATTCTCTTCATGGTTCAGTAGATGAGAAA
CATGGAGATCTTACTTCTAATGGCAGCAATATCAAAGATAGGAAAGGACATTTTGGTCCAGTGATT
AGTCAGAAGCTACATGAACATTTTAAGACACAACCATACCCTAGTCGTTCTCTTAAAGAAAGCTTG
GCAGAAGAACTAGGGCTAACATTTCATCAGGTCAACAGATGGTTTGAAAATAGGCGTCATTTTGCA
AGGTTAGCTTCTTCCAGGAAAGGCATCAGTCCAGATAAGCATAGCCCCCAAAATACTAATAGCCCA
GTGACACCTAGTATGCAACCTAAAGAACCAGAGGGGACTGTGATGGAAGAATCTAATGTATGCATA
AACAGGGATGCTACCATCTCTAAAAAAGCGGTTTCTTCAAAAGTCGGATCAAGAAAGAACCTTAGC
AAGAGTTCCCCTAAAAGCCGGGTCAAAAGTTGA

SEQ ID NO: 208, sequência de polipeptídeo traduzida de PHDf_HD_HOX1b de Zea mays

MEKKTVHVLAEGNGEIENGTSSSQNPETLEHPVLLSASQTVPNNLGIRKNYKRAANRGKKGSQGLT
GQAYTLRSSDIDVRVLRSTSGSKTTSTEHVQPPGQPAAKRRKRSRASNKNSTDEFSQIRKRVRYIL
NRMNYEQSLIEAYASEGWKNQSLDKIRPEKELERAKSEILRCKLRIREVFQNIDSLLSKGKIDESL
FDSEGEISCEDIFCATCGSKNATLGNDIILCDGACDRGFHQNCLNPPLRTEDIPMGDEGWLCPACD
CKLDCIDLINELQGSDISIEDPWEKVFPEAAAMTNGSKQDERFDLPSDDSDDNDFDPNMPEEHVAS
KEEGSSEEEEDDDGGSDSDDSDFLTCSDDLEPLIDKKKVDDLGLSSEDSEDDDYDPAGPDSDKDVE
KKSNSDESDFTSDSDDFCKEIKKSGGHDEVSSPPLPDVKVGDMEKNTAQSNTTSSADDPMETEIDQ
SVVLPVSRRRQAERLDYKRLYDEAYGEASSDSSDEEEWSGKNTPIKSNEEGEVGSPAGKGSRVAHH
NELTTQNTKESLHSLHGSVDEKHGDLTSNGSNIKDRKGHFGPVISQKLHEHFKTQPYPSRSLKESL
AEELGLTFHQVNRWFENRRHFARLASSRKGISPDKHSPQNTNSPVTPSMQPKEPEGTVMEESNVCI
NRDATISKKAVSSKVGSRKNLSKSSPKSRVKS

SEQ ID NO: 209, sequência de ácido nucléico de PHDf_HD_HOX2a de Zea mays X89760

ATGGGAAGGAAAGGGTCAGAAGTATCACCAAGCAAAAGGTATCCCTTGAGGTCTGCACACAGTAGT
GGTAGAGTTCTTCGCTCTGCCTTAGCCAATAATAATAAGGACAGTGAGCCTCTGAATGCGGCTGCT
GCTACCCAAGCAGCTGTGAAAAAAAGAAGTGGCAACCAGCTAGATAGTCCTAATAGCACTGTCAGG

FIGURA 19 CONT.

CTACTTCGTTCTACGTCAAAAAATAAGGACGAGGCACATAGCGGGCCTCTTAATGACAGGATCATT
GATAAGCCAGCTGCAAATAAAAGAAAACATGCCAATCCCTCACAGGTTGGAAGTCCCAGCAGCAGT
GTCAGGGTGCTTCGCTCTGCAATAAAAAATAAAGATATGGCGCGCAGTGAGCCTCAAAGTGACAGG
ACCGCTGGTGAGGCAGCTACAAATAAAAGAAAAAACGCCAATTCCTCAAAGGTGGGGAGTCTGAGC
TCTGGTCTCAGTGTGCTTCGGTCTGCTTCAAAAGCTAAGGATGAGACATGTATCAAACCTTTAAAT
GATAGTCCTGCTGTTGAACCAGCTTCAAGTAAAAGAAAATGCACCAGTCCGTCAAAGGTGGGAAGT
CCCAGCAACAGTGTCAGGGTGCTTCGATCTACTTCAAAATATAAGAATGAGACATGTACCGAGCCT
TTAAATGATAGTGGTGCTGTTGAACCAGCTGCAAATAGAAGGAAAGGTGTCAGTCGTTCAAAGGAG
GGAAGTCCCAATAGCAGTGTCAGGGTGCTTCGCTCTGCCTCAAATAATAAGATTGATGTATCTATT
GAGCCACTGAATAATAGTACTACTGGTCAACCGGCTGCGAAAAGAAGAAAAAATGGCAGTTCCTTT
AAGAGCAGTGCCAGGGTGCTTTGCTCTACCTCAGAAAGGAAAAATGAGGCATCTAGTGAGCCTTTA
AATGATAGTACTGCTGCTCAGCCAGCTGCGAGGAAAAAGAAAGGTGGTGCAATGCCAAAGACTGAT
AACCCAAAGATTGGTGTCAGGGTTCTTCGCTCTGCCTCAGGAAAGAAGAATGAGACATGTACTGAG
CCTGTAAATCGGAGTACTTCTCGTGAACCAACTGTGACAAAAAAAAAGAAATTGCAAGCCTTCAAAG
GATCGCGGTCCCAAGAAGGACTACTTAAAAATTTGTCAAAGAATTAAATATATTTTGAATCGAATG
AAGTATGAACAAGCCTTTATTCAGGCTTATGCAAGTGAAGGTTGGAAAGGCCAAAGTTTGGAAAAA
ATAAGACCCGAGAAGGAGCTTGAACGAGCCAAGGCAGAAATCCTTCAATGCAAGTTAAGAATCCGG
GAAGCTTTTCGAAATATGGACTCTCTTTTATCGAAGGGGAAGCTTGAAGAATCTTTGTTTGATTCA
GCAGGACAAATTTCAAGCGAAGATATATTTTGTGCCATATGTGGCTCAAAGGATGTTACTTCACAA
AATGACATCATTCTTTGTGATGGAGCCTGTGACAGAGGATTCCACCAGAATTGTTTGAGCCCTCCT
TTGCTAACTGAAGAAATTCCTCCGGGGGATGAAGGATGGCTTTGCCCTGCATGTGTGTGCAAAGCA
GACTATATAGATGCGTTAAATGAACTTCAAGGAAGCAAACTCTCCATTCATGACTCATGGGTGAAA
GTTTTCCCTGAGGCAGCTTCAACTGCCAATGGTTCTAAACAAGTTGATGCTTCTGATCTGCTGCCA
GATCATATAAAGGACAGTGCTAACCTTGCATTGGTTGGAACACACATGGTGAATGAAATCAGGTTT
TCTGAAGAGGATGATAGCAAAGCGGATGACCTCGGCTTGCCTTCGGAGGACTCAGGGGATGGTGAC
TTCGATCCAGCAGGTCCAGATTCTAGTGAAGACCAAAATGATGGATTGAACTCAGAAGAGTCAGAT
TTCACATCTGACTCTGATTATTTTTGTGCTGAGATTGCTAAATCTTGCGGCCAGGATGAAGTCTCG
GCATCTCCATTATCAAATGTGATAAACCGTACTGATAGAATGAAACTCAGGGCCTACAGTAGACGC
TCTAATGAAGAAAACCACAATCATGCATTTATGGACATGGAGCTAGAGCAAGACATAGTGCTACCA
GTTCCAAGCAGGCGGCAGGTTGAACGGTTGGATTACAAAAAACTATACGATGAGGCATATGAGGAA
GAATCATCTAATTCAAGTGATGAAGAAGAGTGGTCTGGAAAAGAACTCTTAGAAGGCAGTGGAACA
GATTCACTTGATGAGCCTTCGTCCTGTAAAAGGCTCTCTAGAAGAGCACCAGCAGGGCAGCAGAAT
AATGAACTTACACCACAGAGCCGGCCTCATGGTTCTAACAGTGAGCAGAAAACTGAAGTTCTTCGC
TCCAATGGCAGTAGTAGCACAGGACGGAAGTATGATCCTGAAGTTACCCAGAAATTGAAGGTGCAT
TTTGAGAAAGATCCATATCCTAGCCGTGAAACAAAAGAAAATCTATCAGAAGAACTAGGCTTGACA
TTTAATCAGGTCTCCAAATGGTTCTCTAGTACTCGTCACTATTCAAGGGTTGCTTCTGCCAAAAAA
GAAATGCATCCTGACAGCCATACTTGTGAGAATAATGATGAGACAACTGTAGACAGCATGCAAGCA
AAACAACCCAATGCTGGGGTGATGGAAAAGTTGACGGGGGATAGAAATGGCATTGTTCCTGAGAAA
CCGATGGTGCAGGACAATCTTAACCAAGGAGCGATGGAAAAGTTAACTAGGGATAAAAATGACATT
GTTCCTGAGAAACCGGTGATGCAGAGCAGTCTTAACCAATGTAATAATGAAGATATACCGCTTTCT
GGAACAGAAATTGAGATGGAGTCTTACGAACAAGAATCATCTGACTCAAGTGATGAAGAGTGGTCC
ATACTTAGCACACCACGAAAAACAAGATTACAAGACCATGGAAAATCATCAGTTACTGAATCACTT
CGTTCTGCAAAACGGTGTTCCAGAAGAGCACCAGCTAGGGAGCAGAATAATGAACGTATTCAGAGT
GAGCAGCTTCATGGTTCTGCAAGTGAGCAGCAAACTGAATTTCTTCGCTCCAATGTCAGCAGTGGC
AAGGTCTCAAAATATCATTTTGGCCCTATAGTTAGTCAGAAGCTTAAGGAACATTTTGAAAAAGAT
CCATATCCTTGCCGTGCAACAAAAGAGAGTCTGGCACAAGAGCTAGGACTGACATTTAATCAGATC
AGTAAGTGGTTCTCTGGCACTCGTCATTATTCAAGAGATGCTGTTGCCAAGAATCAGAAACGCCCG
GGAGAAAATACTACTGTGAATAATAATAGCACAACCTTTGATGACATACAAGTAATAGAACCCAAT

FIGURA 19 CONT.

GTTGGGTTGATAGATAAACCAGCTGCAGATGTAAATGACATGATTTCTGAAAAGTTGATGGTCCAA
ATAAATCTTAACGAAGGCATTGAAGAAGATATCCCACCGAGCCAATATGCCAGGTGTGAAGAGAAA
ATAACTATGACTCCAGCTGCCATCTCAAGGGAAGCTGGTCTTCCAGGATATGGTCCTGGTCCTGGA
GAAAACTTTCTTCAGGTCAGTTCAAGGAATACTAGCTGTGAGCAGAGTGTGATACTGGCTCCTACC
GCTATCGCAAGAGAAGTTGGTCCACCAGGATATGCGACTGGAGAAAACCAGGGCAATGGGGCTCCA
TGGAATACAAGTTATGAGCGAAGACTATTTATGAATCCTGTGACCAGCTCAAGAGCAGTCAGTCCT
CCAGGATATGGGCCTGAAGAAAGCCGAGGCAGTGGCATTTCCTGGAATACAAGCAGTGAACAAATA
TTGTTTATGAGTCCTACAACCATCTCAAGAGAACTTGGTCCACCAGGATATGGGCCTAGAGAAGAC
CAGTGCAGTGGCACTTCATGGAATGCAAGCTGCGAGTTGGGAGTATTTATGAGTCCTGTGACCACC
TCAAGAGAAGACAGTCCGCCAGGATATGGGCCTCGAGAAAACCAGGAGAATGACAATATGAGATGC
GAGTTGAGAATGTTTACAAGTCCTACAACCATCTCCAGAGAAGTTGGCCCACCAGGATATGAGCTT
ACTGGAAAAAACAAGGGCAACGGCGCTTCATGTACTACGAATTACCAGCAGGGAGCGTTTACGAGT
CCTGAAGCCTTCCCAAGAGAAGCCTGTCCTCCAGGATTTGTGTCTGGAGAGAACCAGGGTAACGAC
ACTCCATGGTATATGAGCAGCAAACAGATGGCGTTTACGAGTCCTACGACCAACCCTAGAGTAGGT
CCGCCAGGGTATGGACAGGAAAACCAGGGCAGTCGTGTTTCTTGGAATACGAGTTATGAGCAGGGA
GTGTTTATGAGTTCTACAAACATCTCAAGAGAAGCCGGTCCGCCAGGTTATGGGCCTGGGGAAAGT
CAGGGCAATATCACTTCATGGATTTCGACAGGCGGGCAGAGAATGTTTACAGGTCCTTCAACTATC
CCGAACGAAGTCGGGCAGCAAGGATATGGCCCTTCAGGAATCCCGGGAACGGATGGCTCAAGGAGT
ATGGACTTGAAAGCCTTTCCACCAGGATATGGACCTGCTGGAGAAAACCAGGGCGGTGGTGCTTCA
GGGAGCGTCATAAGCCCCCATGCTCGGTCTGCAGAGAACGTAGAGTTTTCAGACGACGCTAGAAAA
AAGGCTATACAGCGGGAGCTGAGACGGCGGCAGAAGTTCAGGTGA

SEQ ID NO: 210, sequência de polipeptídeo traduzida de PHDf_HD_HOX2a de Zea mays

MGRKGSEVSPSKRYPLRSAHSSGRVLRSALANNNKDSEPLNAAAATQAAVKKRSGNQLDSPNSTVR
LLRSTSKNKDEAHSGPLNDRIIDKPAANKRKHANPSQVGSPSSSVRVLRSAIKNKDMARSEPQSDR
TAGEAATNKRKNANSSKVGSLSSGLSVLRSASKAKDETCIKPLNDSPAVEPASSKRKCTSPSKVGS
PSNSVRVLRSTSKYKNETCTEPLNDSGAVEPAANRRKGVSRSKEGSPNSSVRVLRSASNNKIDVSI
EPLNNSTTGQPAAKRRKNGSSFKSSARVLCSTSERKNEASSEPLNDSTAAQPAARKKKGGAMPKTD
NPKIGVRVLRSASGKKNETCTEPVNRSTSREPTVTKKRNCKPSKDRGPKKDYLKICQRIKYILNRM
KYEQAFIQAYASEGWKGQSLEKIRPEKELERAKAEILQCKLRIREAFRNMDSLLSKGKLEESLFDS
AGQISSEDIFCAICGSKDVTSQNDIILCDGACDRGFHQNCLSPPLLTEEIPPGDEGWLCPACVCKA
DYIDALNELQGSKLSIHDSWVKVFPEAASTANGSKQVDASDLLPDHIKDSANLALVGTHMVNEIRF
SEEDDSKADDLGLPSEDSGDGDFDPAGPDSSEDQNDGLNSEESDFTSDSDYFCAEIAKSCGQDEVS
ASPLSNVINRTDRMKLRAYSRRSNEENHNHAFMDMELEQDIVLPVPSRRQVERLDYKKLYDEAYEE
ESSNSSDEEEWSGKELLEGSGTDSLDEPSSCKRLSRRAPAGQQNNELTPQSRPHGSNSEQKTEVLR
SNGSSSTGRKYDPEVTQKLKVHFEKDPYPSRETKENLSEELGLTFNQVSKWFSSTRHYSRVASAKK
EMHPDSHTCENNDETTVDSMQAKQPNAGVMEKLTGDRNGIVPEKPMVQDNLNQGAMEKLTRDKNDI
VPEKPVMQSSLNQCNNEDIPLSGTEIEMESYEQESSDSSDEEWSILSTPRKTRLQDHGKSSVTESL
RSAKRCSRRAPAREQNNERIQSEQLHGSASEQQTEFLRSNVSSGKVSKYHFGPIVSQKLKEHFEKD
PYPCRATKESLAQELGLTFNQISKWFSGTRHYSRDAVAKNQKRPGENTTVNNNSTTFDDIQVIEPN
VGLIDKPAADVNDMISEKLMVQINLNEGIEEDIPPSQYARCEEKITMTPAAISREAGLPGYGPGPG
ENFLQVSSRNTSCEQSVILAPTAIAREVGPPGYATGENQGNGAPWNTSYERRLFMNPVTSSRAVSP
PGYGPEESRGSGISWNTSSEQILFMSPTTISRELGPPGYGPREDQCSGTSWNASCELGVFMSPVTT
SREDSPPGYGPRENQENDNMRCELRMFTSPTTISREVGPPGYELTGKNKGNGASCTTNYQQGAFTS
PEAFPREACPPGFVSGENQGNDTPWYMSSKQMAFTSPTTNPRVGPPGYGQENQGSRVSWNTSYEQG
VFMSSTNISREAGPPGYGPGESQGNITSWISTGGQRMFTGPSTIPNEVGQQGYGPSGIPGTDGSRS
MDLKAFPPGYGPAGENQGGGASGSVISPHARSAENVEFSDDARKKAIQRELRRRQKFR

FIGURA 19 CONT.

SEQ ID NO: 211, sequência de ácido nucléico de PHDf_HD_HOX2b de Zea mays X89761

ATGGGAAGGAAAGGCTCAGAAATGTCACCAAGTAAAAGGTATCCCTTGAGGTCTCCACATAGTAGT
GGTAGAGTTCTTCGCTCTGCCTTAGCCAATGATAATAAGGACAGTGAGCCTCTGAATGCTGCTGCT
GCTACCCAAGCCGCTGTGAAAAAAAGAAGTGGCAGTCAGTTGGATAGTCCTAATAGCAGTGTCAGG
TTGCTTCACTCTACATCAAAAAATAAGAACGAGGTATGTAGCGGGCCTCTAAATGACAGGATCCCT
GATAAGCCAGCTGCAAATAAAAGAAAACGTGCCAGTCCTTCACAGGTGGGAAGTTCCAGCAACAGT
GTCAGGGTCGTTCGCTCTGCAATAAAAAATAAAGATGAGGCATGCAGTAAATCTCTAAGTGACAGG
ACCGCTGGTGAGCCAGCTGCAAATAAAAGAAAAAACGTCAATCGCTCAAAGACGGGGAGTCTGAGC
TGTGGTGTCAGGGTGCTTCTATCTGCTTCAAAAGCTAAGGATGAGACATCTACCAAGCCTTTAAAT
GATAGTGCTACTGTTGAACCAGCTTCAGGTAAAAGGAAAGGCACCAGTCCGTCAAAGGTGGGAAGT
GCCAGCAATAGTGTCAGGGTGCTTCGATCTACTTCAAAATATAAGAATGAGACATGTACTGAGCCT
TTAAATGATAGTGGTGCTGTTGAACCAGCTGCAAATAAAAGGAAAGGTGTCAGTCGTTCAAAGGAG
AGAAGTCTCAATAGCAGTGTCAGGGTCGTTGCTCTGGCCTCAAATAATAAGATTGAGGTATCTATT
AAGCCTCTAGGTAATAGTACTGCTGGGGAACCAGCTGCGAAAAGAAGAAAAAGCGGCAGTTCCTTC
GAGGCAGGAAGCCCTGTGAGCAGTGCCAGGGTGCTTCGCCCTACCTCAGAAAGGAAGAATGAGGCA
TCTAAGCCTTTAAATGAAAGTACTGCTGCTCAACCAGCTGCGAGGAAAAAAAAAAGCTGGGGTAATT
TCAAAGACCGACAATCCAAAGATTGGTCTCAGGGTTCTTCGCTCTGCCTCAGGAAAGAAGAATGAA
GCATGTATTGGGCATGTAAATGATAGTACTTCTGCTGAACCAACTGTGACAAAAAGAAACCGCAAG
CCTTCAATGGATCGCAGTCCCAAGAAGGACTACTTAAAAATTTGTCAAAGAGTTAGATATATTTTG
AATCGAATGAACTATCAACAAACCTTTATTCAAGCTTATGCAAGTGAAGGTTGGAAAGGCCAAAGT
TTGGAAAAAATAAGACCTGAAAAGGAGCTTGAACGAGCCAAGGCAGAAATCCTCCAATGCAAGTTA
AGAATCCGTGAAGCTTTTCGAAATATGGACTCTCTTTTATCCAAGGGGAAGCTTGAAGAATCTTTG
TTTGATTCCGCAGGAGAAATTTCAAGCGAAGATATATTTTGTGCTGTATGTGGCTCAAAGGATGTT
ACTTTACAAAATGACATCATTCTTTGTGATGGAGCCTGTGACAGAGGATTCCACCAGAACTGTTTG
AACCCTCCTTTGCTAACTGAAGATATACCTCCGGGGGATCAAAGATGGCTTTGCCCTGCATGTGTC
TGCAAAGCAGACTCTATAGATGCACTAAATGAACTTCAAGGGAGCAAGCTCTCCATTCATGACTCA
TGGGAGAAAGTTTTCCCTGAGGCAGCTTCAATTGCCAACGGTTCTAAACAAGTTGATACTTCTGAT
CTGCTGCCAGATCATATAAAGCACAGTGACAACCCTGCATTGGTTGAAGGGCTGATGGTGGATGAA
GTCAGGCTTTCTGCAGAGGATGACAGCAAAGCGGATGACCTTCGCTTGTCTTCGGAGGACTCAGGG
GATGGTGACTTCGATCCGTCAGGTCCAGATTCTAGTGAAGACCAAAAAGATGGATTGAACTCAGAA
GAGTCAGATTTCACATCAGACTCTGATGATTTTTGTGCTGAGATTGCCAAATCTTGTGGCCAGGAT
GAAGTCTCGGCATCTCCATTATCAAATGTGATAAACCATACTTATAGAATGAAACTCAGGGCCAGC
AACAACCGCTCTAATGAAGAAAACCATGATCATGTATTTATGGACATGGAGCTAGGGCAAGACATG
GTGCTACCAGTTTCAAGCAGGCGGCAGGTGGAACGGTTGGATTACAAAAAACTATATGATGAGGCA
TATGGGAAAGAATCATCTAATTCAAGTGATGACAAAGAGTGGTCTGGAAAAGAACTATTAGAAGGC
AGTGAAACAGATTCACTTAGTGAGCGACCTCATCCTGTAAAACGGTGCTCTAGAAGAGCACAAGCA
GAGCAGCAGAATAATGAACATACACCACAGAGGGAGCGCCTTCATGGTTCTGAAAGTGAGCAGAAA
ACTGGAATTCTTCGATCCAATGGCAGCAGTAGCACAGGACGGAAGTTTGGTCCTGTAGCTACCCAG
AAATTGAAGGTACATTTTGAGAAAGATCCATATCCTAGCCGTGAAACGAAAGAGAATATATCAGAA
GAACTAGGCTTGACATTTAATCAGGTCTCCAGATGGTTCTCTAGCACTCGTCACTATTCAAGGGTT
GCTTCTGCCAGAAAAGAAATGCATCCTGACGGCCATACTAGCGAGAATAATGATACCACGACTGTA
GATAGCATGCAAGCAAGGCAACCCAACACTGTGGTGATGGAAAAGTTGACTGGGAATAGAAATGAC
ATTGTTTCTGAGAAACCGATGGTGCAGAACAATCTTAACCAAGGGGTGGCAGAAAAGTTGACTAGG
GATAGAAATGACATTGTTCCTGAGAAACAGGTGGTGCAGAGCAATCTTAACCAATGTAATAATGAA
GATATCCCGCTTTCTGGAACAGAAATTGAGATGGAGTCTTACGAACAAGAATCATCTGAAAGCAGT
GATGAAGAGTGGTCCACACTTAGCACACCACGAAAAACAAAATTACAAGGCAACGAAAAAGTATCA
CTTACTGAATCACTACGTTCTGCAAAACGGTGTTCTAGAAGAGCACCAGCTAGGGAGCAGAATAAT

FIGURA 19 CONT.

GAACATACTCAGAGTGAGCAACTTCATGGTTCTGTAAGTAAACAGCAAACTGAAGTTCTTCGCTCC
AATGTCAGCAGTGGCGATGCCTCAAAATGTCATTTTGGCCCTATAGTTACTCAGAAGCTGAAGGAA
CATTTTGAAAAAGATCCATATCCTTGCCGTGCAACAAAAGAGGGTCTGGCACAAGAGCTAGGACTG
ACATTTAATCAGATCAGTAAGTGGTTCTCTGCCACCCATCATTATTCAAGAGATGCTGTTGCCAAG
AACCAAAAATACCCAGGAGAAAATACTACTGAGAATAATAGTAGCACAATCTTTGATGGTATACAA
GTAATAGAACCCAATTTTGGTTTGATAGATAAACCAGATGCAGATACAAATGACATGATTTCTGAA
AAGTTGATGGTACAAATTAATCTTAACGAAGGCATTGAAGAGGATATCCCACCGAGCCAATATACT
ACCAGGTGTGAAGAGAAATTAACTATGACTCAAGCTGCCATCTCAAGGGAAGCTGGACCTCCAGGT
TATGGTCCTGGAGAAAACTTTCTTCAGGTCAGTTCAAGGATTACTAGCTGTGAGCAGAGTGTGATA
ATGGCTCCTAGCGCTATCGCAAGAGGAGTTGGTCCACCAGGATATACGCCTGGAGAACACCAGGGC
AATGGGGCTCCATGGAATACAAGCTATGAGCGAAGACCGTTTACGAGTCCTGCGACTAGCTCAAGA
GAAGTCGGACCTCCAGGATATGGGCCTGAAGAAAACCAAGGCAGTGGCATTTCCTGTAATACAAGC
AGAGAGCAAATATTGTTTATGAGCCCTACAACCATCTCGAGAGAACTTGGTCCACCAGGATATTTG
CCCAGAGAAAACCAGCGCAGTGACACTTCATGGAATACAAGCTGCGAGTCGGGACTATTTATGAGT
CCTGCGACCACCTCAATAGAACATGGTCCACCAGGATATGGGCCTGGAGAAAACCAGGAGAATGAC
AGTACAAGCTGCGAGTTGAGAATGTTTACCAGTCCTACAATCGTCTCCAGAGAAGTTGGCCCACCA
GGATACCAGCAGGGAGTGTTTATGAGTCCTAAAACCATCTCAAGAGAAGTCTGTCCTCCAGGATTT
GGGTCTGGAGAGAACCAGGGCAACGACACTACATGGTACAAACAGAGCGCGTTTACAAGTCCTACA
ACCATCTCGAGAGAAGTCGGTCCGCCAGGGTATGGACAGGAAGATCAGGGCAATGGTGGTTCTTGG
AACACGAGTTGTGAGCAGGGAGTGTTTACGAGTCCTACTAGAGGCATCTCAAGAGAAGTCGGTCCG
CCAGGATACGGGCCCGGGGAAAGTCAGGGCAATACCACTTCGTGGATTTCGACCTGCCAGCAGAGA
ATGCTTGCAGGAATCCAGAGAACTGGTGGCTCAAGGAGTATGGACTTGGAAGCCTTTCCACCGGGA
TACGGACCTGGTGGTGCTTCAGGGAGTGTCATAAGCCCCCAAGCTCGGTCTGCAGAGAATGTAGAA
TTTTCGGACGAGGCTAGAAAAAAGGCTATACAGCGGGAGCTGAGACGGCGGCAGAAGTTCAGGTGA

SEQ ID NO: 212, sequência de polipeptídeo traduzida de PHDf_HD_HOX2b de Zea mays

MGRKGSEMSPSKRYPLRSPHSSGRVLRSALANDNKDSEPLNAAAATQAAVKKRSGSQLDSPNSSVR
LLHSTSKNKNEVCSGPLNDRIPDKPAANKRKRASPSQVGSSSNSVRVVRSAIKNKDEACSKSLSDR
TAGEPAANKRKNVNRSKTGSLSCGVRVLLSASKAKDETSTKPLNDSATVEPASGKRKGTSPSKVGS
ASNSVRVLRSTSKYKNETCTEPLNDSGAVEPAANKRKGVSRSKERSLNSSVRVVALASNNKIEVSI
KPLGNSTAGEPAAKRRKSGSSFEAGSPVSSARVLRPTSERKNEASKPLNESTAAQPAARKKKAGVI
SKTDNPKIGLRVLRSASGKKNEACIGHVNDSTSAEPTVTKRNRKPSMDRSPKKDYLKICQRVRYIL
NRMNYQQTFIQAYASEGWKGQSLEKIRPEKELERAKAEILQCKLRIREAFRNMDSLLSKGKLEESL
FDSAGEISSEDIFCAVCGSKDVTLQNDIILCDGACDRGFHQNCLNPPLLTEDIPPGDQRWLCPACV
CKADSIDALNELQGSKLSIHDSWEKVFPEAASIANGSKQVDTSDLLPDHIKHSDNPALVEGLMVDE
VRLSAEDDSKADDLRLSSEDSGDGDFDPSGPDSSEDQKDGLNSEESDFTSDSDDFCAEIAKSCGQD
EVSASPLSNVINHTYRMKLRASNNRSNEENHDHVFMDMELGQDMVLPVSSRRQVERLDYKKLYDEA
YGKESSNSSDDKEWSGKELLEGSETDSLSERPHPVKRCSRRAQAEQQNNEHTPQRERLHGSESEQK
TGILRSNGSSSTGRKFGPVATQKLKVHFEKDPYPSRETKENISEELGLTFNQVSRWFSSTRHYSRV
ASARKEMHPDGHTSENNDTTTVDSMQARQPNTVVMEKLTGNRNDIVSEKPMVQNNLNQGVAEKLTR
DRNDIVPEKQVVQSNLNQCNNEDIPLSGTEIEMESYEQESSESSDEEWSTLSTPRKTKLQGNEKVS
LTESLRSAKRCSRRAPAREQNNEHTQSEQLHGSVSKQQTEVLRSNVSSGDASKCHFGPIVTQKLKE
HFEKDPYPCRATKEGLAQELGLTFNQISKWFSATHHYSRDAVAKNQKYPGENTTENNSSTIFDGIQ
VIEPNFGLIDKPDADTNDMISEKLMVQINLNEGIEEDIPPSQYTTRCEEKLTMTQAAISREAGPPG
YGPGENFLQVSSRITSCEQSVIMAPSAIARGVGPPGYTPGEHQGNGAPWNTSYERRPFTSPATSSR
EVGPPGYGPEENQGSGISCNTSREQILFMSPTTISRELGPPGYLPRENQRSDTSWNTSCESGLFMS

FIGURA 19 CONT.

PATTSIEHGPPGYGPGENQENDSTSCELRMFTSPTIVSREVGPPGYQQGVFMSPKTISREVCPPGF
GSGENQGNDTTWYKQSAFTSPTTISREVGPPGYGQEDQGNGGSWNTSCEQGVFTSPTRGISREVGP
PGYGPGESQGNTTSWISTCQQRMLAGIQRTGGSRSMDLEAFPPGYGPGGASGSVISPQARSAENVE
FSDEARKKAIQRELRRRQKFR

SEQ ID NO: 213, sequência de ácido nucléico de PHDf_HD_II de Vitis vinifera contíguo de AM464161.2, AM478203.2

ATGGATGCTTCCCCTGCCCAAGAGAGTAATCGGACTAGGAAAAGTTCTAGTCCTAAACAGAATATT
TTAGAAGAGGCCAGAAAACTTTCTGAAAGTGTGTGTAGCGAATCGTCAGAGCAGAAACGTCCCTCT
GAAAATGGACAGCATGAGCCAGCAGAAATCAGCCCTGTGCTATCCAATTGTATTGTTACTGAACAG
TCAGAACTGCCTCCAGAAGATGTGGGTGATACAATATTAGGATTGCCTCCAGCAGATGTGACCAAG
AATTCTCTTACTGAACATTTGGGATTGCCTCCAGAAGATGCAATCAAGAATGATGGAACTGAACAA
TTGGGGTTTTTTCCTGAAGTTGTGACCAAGAGTTCTATTATTGAAAAATTAGGACAATCAGAGCCA
CCTCCTGAAAATGTGGCCCGGTATTCTGGTCTTGACCAGTCAGGGTCAGCACCCAAAGATTTGGCC
AATAAAAGAACTGCAAAACTAGTTAAAAGAAAATATAAATTGAGATCTTCAGTAAGCGGTTCCCGG
GTTTTACGCTCAAGGTCACAAGAGAAACCTAAAGCTTCACAGCCAAGTGATAATTTTGTAAATGCC
AGTGCTAGCAGAGAAAGGAAAGGAAGAAAGAAGAAAAGGATGAATAAAACAACTGCAGATGAATTT
GCGAGAATCAGGAAACATCTTAGATATTTACTGAACAGAATGAGCTATGAGCAAAATCTGATTGAT
GCTTATTCTGCCGAAGGTTGGAAAGGACAAAGTGTGGAAAAATTAAAGCCAGAGAAGGAACTTCAA
CGTGCCTCATCTGAAATTTCTCGTCGCAAACTGMAAATACGTGATCTATTTCAACATCTTGATTCA
TTATGTGCTGAAGGAAGGTTTCCAGAATCTTTATTTGATTCTGAAGGGCAGATTGACAGTGAGGAT
ATATTCTGTGCTAAATGTGAGTCCAAAGACATGTCTGCTGATAATGACATAATACTCTGTGACGGT
GCTTGTGATCGTGGATTTCACCAGTTTTGTCTGGAACCGCCATTGTTAAAAGAAGAAATTCCTCCT
GATGACGAGGGTTGGCTGTGCCCTGCATGTGATTGCAAAGTTGACTGCATGGACCTGCTTAATGAC
TCTCAAGGAACAAAACTTTCTGTAATTGATAGCTGGGAGAAGGTTTTTCCTGAGGCAGCTGCAGCT
GGGAATAACCAGGATAACAACTCTGGATTTTCATCAGATGATTCTGAGGATAATGATTATGACCCC
GATTGTCCAGAGGTTGATGAGAAGGGTCAGGGAGATAAGTCAAGTTCTGATAAATTTGATGAATCT
GATGAATTTGACGAATCTGATGAATCTGATTTCACTTCTGCATCTGATGATATGGTGGTCTCACCR
AATAATGAGCAGTGTTTGGGGCTTCCTTCTGATGATTCAGAGGATGATGATTTTGATCCTGATGCT
CCAGAAATTGATGAACAGGTTAATCAGGGGAGTTCAAGTTCTGATTTTACATCTGACTCCGAGGAT
TTTACTGCCACTTTAGATCGCAGAAACTTCTCTGACAACGAGGATGGTCTTGATGAGCAAAGAAGA
TTTGGTAGGAAGAAGAAAGATACTTTAAAGGATGAGCTCTTATCCGTGTTAGAGTCAAATTCTGGT
CAAGATAATGCACCTCTGTCTGCAAAGAGACATGTAGAAAGGCTGGATTACAAAAAGCTGCATGAT
GAGGCATATGGAAATGTTTCTTCTGATTCAAGTGATGATGAAGACTGGACAGAGAATGTTATACCA
AGAAAGAGGAAGAATCTTAGTGGTAATGTTGCCTCAGTGTCACCAAATGGAAACACTTCAATCACT
GAGAATGGGACAAATACCAAGGACATAAAGCACGACTTGGAAGCAGCTGGATGCACTCCCAAGCGA
AGAACTCGTCAGAAGCTGAACTTTGAAAGCACAAATAATTCACTTGCTGAGTCACATAAAGGTTCT
CGAAGTCCTGGTTCTACTGGTGAAAAAAGTGGGCAATCATCATATAAAAAACTTGGAGAAGCTGTA
ACTGAGAGACTCTACAAATCCTTCCAGGAAAATCAGTACCCTGATCGTGCTATGAAAGAAAAATTG
GCAGAAGAGCTGGGAATAACGAGCCGGCAGGTTAGCAAATGGTTTGAGAATGCCCGTTGGAGCTTT
CGCCATCGACCACCCAAGGAAGCCAGTGCGGGTAAAAGTGCTGTGAAAAAGGATGCATCCACGTCT
CAAACAGATCAAAAGCCTGAGCAAGAAGTGGTTCTTAGAGAAAGTTCTCACAATGGAGTGGGAAAA
AAGGAGTCGCCCAAAGCAGGTGCTTCGAAGGTAGACCGCAGCAAGGAAGCCAATGCAGGTAAAAGT
GCTGTGAAAAAGGATGCATCCACGTCTCAAACAGATCAAAAGCCTGAGCAAGAAGTGGTTATTAAA
GAAAGTTCTCACAATGGAGTGGGAAAAAAGGAGTCGACCAAAGCAGGTGCTTCGAAGGTAGACCGC
TGCAGTGGAGCTAAGAGGAGAAGGAAATTGGCAACTGATGGAAGCCATAGACAGAAATCTTCAACT
CCCAACTCTACAAGACAAAAGACAAAGTCAAATCATGAAGCATCTGAGGCAACCAATGGAAGCAGA
AGGCAGAATTCTTCTACTCCCAAGTCTAGAAGGCAAAAGACTAAGTTAGCTGGTGAAGCGTCTGAA

FIGURA 19 CONT.

GCAACTGGTGGAAGCAGTAGACAGAATTCGTCAACCCCAAATTCTAAAAGGCGAAAAACCAAGTCA
GATCATGAAGCATCTAATCCCGTCCTTAGCGGTAAGAAGATAGCAAAGACGGCAAAGAGCTCATCT
GGCACACCAAAAACAAAGGAAAAGCTAAGTGACAGAATCCAAACAAGAAGTAGGAAATCTATTGCT
TGA

SEQ ID NO: 214, sequência de polipeptídeo traduzida de PHDf_HD_II de Vitis vinifera

MDASPAQESNRTRKSSSPKQNILEEARKLSESVCSESSEQKRPSENGQHEPAEISPVLSNCIVTEQ
SELPPEDVGDTILGLPPADVTKNSLTEHLGLPPEDAIKNDGTEQLGFFPEVVTKSSIIEKLGQSEP
PPENVARYSGLDQSGSAPKDLANKRTAKLVKRKYKLRSSVSGSRVLRSRSQEKPKASQPSDNFVNA
SASRERKGRKKKRMNKTTADEFARIRKHLRYLLNRMSYEQNLIDAYSAEGWKGQSVEKLKPEKELQ
RASSEISRRKLXIRDLFQHLDSLCAEGRFPESLFDSEGQIDSEDIFCAKCESKDMSADNDIILCDG
ACDRGFHQFCLEPPLLKEEIPPDDEGWLCPACDCKVDCMDLLNDSQGTKLSVIDSWEKVFPEAAAA
GNNQDNNSGFSSDDSEDNDYDPDCPEVDEKGQGDKSSSDKFDESDEFDESDESDFTSASDDMVVSP
NNEQCLGLPSDDSEDDDFDPDAPEIDEQVNQGSSSSDFTSDSEDFTATLDRRNFSDNEDGLDEQRR
FGRKKKDTLKDELLSVLESNSGQDNAPLSAKRHVERLDYKKLHDEAYGNVSSDSSDDEDWTENVIP
RKRKNLSGNVASVSPNGNTSITENGTNTKDIKHDLEAAGCTPKRRTRQKLNFESTNNSLAESHKGS
RSPGSTGEKSGQSSYKKLGEAVTERLYKSFQENQYPDRAMKEKLAEELGITSRQVSKWFENARWSF
RHRPPKEASAGKSAVKKDASTSQTDQKPEQEVVLRESSHNGVGKKESPKAGASKVDRSKEANAGKS
AVKKDASTSQTDQKPEQEVVIKESSHNGVGKKESTKAGASKVDRCSGAKRRRKLATDGSHRQKSST
PNSTRQKTKSNHEASEATNGSRRQNSSTPKSRRQKTKLAGEASEATGGSSRQNSSTPNSKRRKTKS
DHEASNPVLSGKKIAKTAKSSSGTPKTKEKLSDRIQTRSRKSIA

SEQ ID NO: 215, sequência de ácido nucléico de PHDf_HD_II de Populus tremuloides scaff_IX. 730

ATGTCTGAGGCTGAGAACATGGGTGTTTCTCCATCACAAGTTAGTTCTCAAACTAAAAGTTACTCA
TGCCCTTCACAAACTAAACTGGAAAATACTCATGGATTTACTGCTGAATACAACTGTGGTGGATAT
TCGGAAGAGAAACATAAGCTTGAGTCTGAAATCATACAAACTGAAGCCGGAGATAGCGGTACTGCA
GTTCTCCAATCTGGTGCTGGGGAAACTGTGGAACCATCTACTGAAGATGTAACAAACAATTCTTTT
ACTGATTTGGATCCACCTCCCGAAGATGCCAGGGGTGCTACTTTTGATGAAAGTTCAAGACCTATA
CTAACTGCTATAGATCAGAAACTCGAGCCAGGTGCTAGAAGTGTGAATACTGCGTGTACGCATGGT
GAATCATCAAAAGCAACTGACTCTGGTATTCTACAGGATGAACCTGGAAATACCAATGCTGCATCA
TCAAGCTGCATTGCTAATGAAACCTCACAAGCATCACTTGAAAATCTAGCCAATAATTCTTGTACT
GAAGATGTGGGTCCGCCGTATGGAGATGCAAGCAAGGGTAACCAGATTGATAAAAGTTCATACCCT
CAACAAACTATATCTGGACATACACTTGAGTTACTTTCTGACAGAGCTTGTTGTGAACGATCAGAA
GAAAGACAGAAGCCTGGATCTGAACTTTCAGAAAATGAATCAACGGGAATTGACACTGAATTATAT
AGTGGCATTGCTATCGAGAATTCAGAGCCGCTTACTCAACTAGTGACCAAGAGTTCTCCTATCAAA
CATGTAGGCTTGCTTCCTGGTGACAGCATCATCATTCCTGCAAATGAACAAACAAGACCAACTCAT
GATGATGAGGACAAAGGCCCAGATCATGAACATTTGGAAACACCATCTAGAGTTGCGATTGGTATT
ACTAGACGCGGTAGACCTAGAGGTAAGAGTGCTTCAAGATTGTCAAGGAAGATATATATGTTAAGA
TCTTTGAGAAGTAGTGATAGAGTTCTCCGGTCAAGATCACAAGAGAAACCTAAAGCTCCTGAATCA
AGTAATAATTCAGGTAATGTTAATTCCACTGGTGATAAAAAAGGGAAAAGAAGGAAAAAGAGAAGA
GGAAAGAATATTGTGGCTGATGAATATTCCAAAATCAGGGCACATCTTAGGTACTTATTGAATCGG
ATGAGCTATGAGCAAAGCTTGATCACTGCTTATTCTGGGGAAGGCTGGAAAGGACTTAGCCTAGAA
AAGTTAAAGCCCGAGAAGGAGTTGCAGCGAGCCACATCTGAAATTACCAGACGCAAAGTGAAAATA
AGGGATTTATTTCAACATATTGATTCTCTGTGCAGCGAAGGAAGGTTCCCATCATCTTTATTTGAT
TCTGAAGGACAGATTGACAGTGAGGATGTATTCTGTGCAAAATGTGGGTCCAAAGATTTGAATGCT
GATAATGATATTATACTATGTGATGGTGCTTGTGACAGAGGATTTCATCAGTTCTGTTTGATACCA

FIGURA 19 CONT.

CCATTGTTAAGAGAAGACATTCCTCCTGATGATGAGGGCTGGTTATGCCCTGGATGTGATTGCAAA
GTTGACTGCATTGGTTTGCTTAACGATTCTCAAGGAACAAATATATCTATCAGTGATAGCTGGGAG
AAAGTTTTCCCTGAGGCTGCCGCCACAGCGTCTGGACAAAAACTGGATCACAACTTTGGACCATCT
TCTGATGATTCTGATGATAATGATTATGAACCTGATGGTCCAGATATTGATAAGAAGAGTCAGGAG
GAGGAATCGAGCTCTGACGAATCTGACTTCACTTCTGCATCTGATGAATTCAAGGCTCCCCCGGAT
GGCAAAGAGTACTTGGGGCTCTCTTCTGATGATTCAGAGGATGATGACTATGATCCTGATGCTCCA
GTTCTTGAGGAGAAGCTGAAGCAGGAAAGTTCAAGTTCTGATTTTACATCAGACTCTGAGGATCTT
GCTGCAACTATTAATGGTGATGGGTTGTCTTTAGAAGATGAATGCCACATGCCTATTGAACCTCGT
GGAGTTTCTAATGGGAGGAAATCTAAATTTGATGGAAAGAAGATGCAGTCTTTAAACAGCGAGCTT
TTATCCATGCTAGAACCAGATCTGTGCCAAGATGAGTCAGCAACTGTTTCTGGAAAAAGAAATGTT
GACAGATTGGACTACAAGAAGCTCTATGATGAAACGTATGGAAACATTTCTACTTCAAGTGATGAT
GATTACACTGATACTGTTGGGCCAAGGAAGAGGAGGAAAAATACTGGAGATGTTGCTACAGTGACA
GCAAACGGAGATGCCTCTGTTACCGAGAATGGAATGAATAGCAAGAATATGAACCAGGAATTAAAG
GAGAACAAACGTAATCCTGAGAGAGGAACTTGCCAAAACTCGAGCTTTCAAGAAACAAATGTTTCC
CCGGCTAAATCATATGTTGGTGCATCTCTATCTGGTTCTAGTGGTAAAAGCGTTAGGCCCTCTGCT
TATAAAAAACTTGGAGAAGCTGTAACTCAGAGACTGTACAGTTACTTCAGGGAAAATCAGTATCCA
GACCGAGCTGCAAAAGCAAGCTTGGCAGAAGAGCTAGGAATTACTTTTGAGCAGGTTAACAAGTGG
TTTGTGAATGCCCGTTGGAGCTTCAACCATTCATCATCTACAGGTACAAGTAAAGCTGAAAGTGCT
TCTGGAAAAGGTAGCTGTGATGGCCAAGTGAGGGATAGTGAATCGAAGAACCGAAAGAGCAATAAA
CAAAAAACTAACACCCCGAAATCTAGGAGATAA

SEQ ID NO: 216, sequência de polipeptídeo traduzida de PHDf_HD_II de Populus tremuloides

MSEAENMGVSPSQVSSQTKSYSCPSQTKLENTHGFTAEYNCGGYSEEKHKLESEIIQTEAGDSGTA
VLQSGAGETVEPSTEDVTNNSFTDLDPPPEDARGATFDESSRPILTAIDQKLEPGARSVNTACTHG
ESSKATDSGILQDEPGNTNAASSSCIANETSQASLENLANNSCTEDVGPPYGDASKGNQIDKSSYP
QQTISGHTLELLSDRACCERSEERQKPGSELSENESTGIDTELYSGIAIENSEPLTQLVTKSSPIK
HVGLLPGDSIIIPANEQTRPTHDDEDKGPDHEHLETPSRVAIGITRRGRPRGKSASRLSRKIYMLR
SLRSSDRVLRSRSQEKPKAPESSNNSGNVNSTGDKKGKRRKKRRGKNIVADEYSKIRAHLRYLLNR
MSYEQSLITAYSGEGWKGLSLEKLKPEKELQRATSEITRRKVKIRDLFQHIDSLCSEGRFPSSLFD
SEGQIDSEDVFCAKCGSKDLNADNDIILCDGACDRGFHQFCLIPPLLREDIPPDDEGWLCPGCDCK
VDCIGLLNDSQGTNISISDSWEKVFPEAAATASGQKLDHNFGPSSDDSDDNDYEPDGPDIDKKSQE
EESSSDESDFTSASDEFKAPPDGKEYLGLSSDDSEDDDYDPDAPVLEEKLKQESSSSDFTSDSEDL
AATINGDGLSLEDECHMPIEPRGVSNGRKSKFDGKKMQSLNSELLSMLEPDLCQDESATVSGKRNV
DRLDYKKLYDETYGNISTSSDDDYTDTVGPRKRRKNTGDVATVTANGDASVTENGMNSKNMNQELK
ENKRNPERGTCQNSSFQETNVSPAKSYVGASLSGSSGKSVRPSAYKKLGEAVTQRLYSYFRENQYP
DRAAKASLAEELGITFEQVNKWFVNARWSFNHSSSTGTSKAESASGKGSCDGQVRDSESKNRKSNK
QKTNTPKSRR

SEQ ID NO: 217, sequência de ácido nucléico de PHDf_HD_III de Populus tremuloides LG_I002624

ATGTCCGAGGCTGAGCACATGGGTGTTTCTCCATCAAAAGTTTCTCATACTAAAAGTTATTCATGC
CCTGCACAAACTACACTGGAAAATACGCATGAACCCAGTGCTGAATACAAGTTTGGTGGATATCCG
GAAGAGAGACATAAGCTTGAGTGTGAAATCATACAAACTGAAGCTGGAGATAACAGGGCTGCAGTT
CTCCAATCTTGTTCTGGTGAAGTTGTGCAACCATCTACTGACGATTTAACAAAAAGTCCTCTTATT
GATTTGGACCCACCTCCTGACGATGCAAGGAGTGCTCTGTTTGATAATAGTCCAAGACCTATATCA
ACCGCTATGGATCAGAAACTTGAGCCAGGTGCTACAAGTGTGAATACTGCTTGTGTGCATAGTGAA
TCATCAAAAGCAATTGACTCTAGTATTCTACTGGATGAACCTAGGAATAGCAACACAGAATTGTCA

FIGURA 19 CONT.

AGCTGCATTGCAAATGAAACCTCACAAGCATCACTTGAAGGTCTAGCCAATGATTCTCGTGCTGAA
GATGCAGGACTGTCACTTGTAGAGGCAAGCAATAGTGACCTGATTGATGAAAGTTCATACTCTCAA
CAAACTACATCTGGACAGACACGTGAGTTTCATTCTGACAGAGCTTGTTGTAAACCATTGGAAGAA
AGACAGAAGCCCGGCTCTGAACTTGCAGAAAATGAATCAATGGAAATTGGCATTGGATTACCCAGT
GGTATTGCTATTGAGAATTTGGAGCCGCTTACTGAACTTGTGACCAAGAGTTGTCCTATCAAACAT
ATAGGCTTGCCTCCTGGTGATGACATCAGCATTCCTGCAAATGAACAAATAAGACCTACTCATGAT
AAGGAGTCCAAATACCCTGATTGTGAACATTTGGAAAAACTATCTGGAATTGTGATTGGTATTACT
AGTCAAGGTGTACCCAGTGTTAAGAGAACTTCAAAATTGTCAGGGAAGAAATATACAAGTTCTTCG
AGAAAAAGTGATAGAGTTCTCCGGTCAAATTCACAAGAGAAACCTAAAGCTCCCGAGCCGAGTAAT
AATTCTACTAATGTTAATTCCACTGGTGAGGAAAAAGGGAAAAGGAGGAAAAAGAGAAGAGGGAAA
AGTATAGTGGCTGATGAATATTCCAGAATCAGGGCACGTCTTAGGTACTTGTTGAATCGGATGAGC
TATGAGCAAAGCTTGATCACTGCTTATTCCGGGGAAGGCTGGAAAGGACTTAGTCTAGAAAAGTTA
AAACCTGAGAAGGAGTTGCAACGGGCCACATCTGAAATCATCAGACGCAAAGTGAAAATAAGGGAT
TTATTTCAACATATTGATTCACTATGCGGTGAAGGAAGGTTCCCAGCATCTTTGTTTGATTCTGAA
GGACAAATTGACAGTGAGGATATATTCTGTGCAAAATGTGGATCCAAAGATTTGACTGCTGATAAT
GATATTATACTATGTGATGGTGCTTGCGACCGAGGGTTTCATCAATTCTGTTTGGTACCACCATTG
TTAAGAGAAGACATTCCTCCTGGTGATGAGGGCTGGTTATGCCCTGGATGTGATTGCAAAGTTGAC
TGCATTGACTTGCTTAATGACTCTCAAGGAACAAATATATCTATCAGTGACAGATGGGATAATGTT
TTCCCCGAGGCAGCTGCAGTAGCATCTGGACAAAAACTCGACTACAACTTTGGATTGTCTTCTGAT
GATTCTGATGATAATGATTATGATCCTGATGGTCCAGATATTGACGAGAAGAGTCAGGAGGAATCA
AGCTCTGACGAATCTGACTTCAGTTCTGCGTCTGATGAATTTGAGGCTCCACCTGATGATAAACAG
TACTTGGGGCTCCCTTCTGATGATTCAGAGGATGATGACTATGATCCTGATGCTCCAGTTCTCGAG
GAGAAGCTGAAGCAGGAAAGTTCCAGTTCTGATTTTACTTCGGACTCTGAGGATCTTGATGCCACT
CTTAATGGTGATGGGTTGTCTTTAGGAGATGAATACCACATGCCTATTGAACCTCATGAAGATTCT
AATGGGCGAAGATCGAGATTTGGTGGAAAGAAGAATCATTCTTTAAACAGCAAGCTTTTATCCATG
CTAGAACCAGATTCTCACCAAGAAAAGTCTGCACCTGTTTCTGGAAAAAGAAATATTGAAAGATTG
GACTACAAGAAGCTCTATGATGAAACATATGGAAACATTTGTACTTCAAGTGATGATGATTTCACT
GATACTGTTGCACCAAGGAAGAGGAGGAAAAACACTGGAGATGTTGCTATGGGGATAGCAAATGGA
GATGCTTCTGTTACTGAGAATGGATTGAATAGCAAGAATATGAATCAGGAATTAAAGAAGAATGAA
CATACTTCTGGGAGAACTCACCAAAACTCGAGCTTTCAAGATACAAATGTTTCCCCAGCAAAAACA
CATGTTGGCGAATCTCTCTCTGGTTCTAGCAGTAAAAGAGTTAGGCCCTCTGCTTATAAAAAACTT
GGAGAAGCTGTAACTCAGAAACTGTACAGTTTCTTCAAGGAAAATCGGTATCCTGACCAAGCTGCA
AAGGCAAGTTTGGCAGAAGAGCTAGGAATTACTTTTGAGCAGGTTAACAAGTGGTTCATGAATGCT
CGTTGGAGCTTCAACCATTCATCACCTGAAGGTACAAGTAAAGCTGAAAGTGCTTCAGGAAAGGGT
AGCTGTGATGGGCATGTGAGGGATAGTGAATCGAAGAACCAAAAGAGCAATAAACAAAAAACTAGT
ACCCCAAAATCTAGGAGATAG

SEQ ID NO: 218, sequência de polipeptídeo traduzida de PHDf_HD_III de Populus tremuloides

MSEAEHMGVSPSKVSHTKSYSCPAQTTLENTHEPSAEYKFGGYPEERHKLECEIIQTEAGDNRAAV
LQSCSGEVVQPSTDDLTKSPLIDLDPPPDDARSALFDNSPRPISTAMDQKLEPGATSVNTACVHSE
SSKAIDSSILLDEPRNSNTELSSCIANETSQASLEGLANDSRAEDAGLSLVEASNSDLIDESSYSQ
QTTSGQTREFHSDRACCKPLEERQKPGSELAENESMEIGIGLPSGIAIENLEPLTELVTKSCPIKH
IGLPPGDDISIPANEQIRPTHDKESKYPDCEHLEKLSGIVIGITSQGVPSVKRTSKLSGKKYTSSS
RKSDRVLRSNSQEKPKAPEPSNNSTNVNSTGEEKGKRRKKRRGKSIVADEYSRIRARLRYLLNRMS
YEQSLITAYSGEGWKGLSLEKLKPEKELQRATSEIIRRKVKIRDLFQHIDSLCGEGRFPASLFDSE
GQIDSEDIFCAKCGSKDLTADNDIILCDGACDRGFHQFCLVPPLLREDIPPGDEGWLCPGCDCKVD
CIDLLNDSQGTNISISDRWDNVFPEAAAVASGQKLDYNFGLSSDDSDDNDYDPDGPDIDEKSQEES

FIGURA 19 CONT.

SSDESDFSSASDEFEAPPDDKQYLGLPSDDSEDDDYDPDAPVLEEKLKQESSSSDFTSDSEDLDAT
LNGDGLSLGDEYHMPIEPHEDSNGRRSRFGGKKNHSLNSKLLSMLEPDSHQEKSAPVSGKRNIERL
DYKKLYDETYGNICTSSDDDFTDTVAPRKRRKNTGDVAMGIANGDASVTENGLNSKNMNQELKKNE
HTSGRTHQNSSFQDTNVSPAKTHVGESLSGSSSKRVRPSAYKKLGEAVTQKLYSFFKENRYPDQAA
KASLAEELGITFEQVNKWFMNARWSFNHSSPEGTSKAESASGKGSCDGHVRDSESKNQKSNKQKTS
TPKSRR

SEQ ID NO: 219, sequência de ácido nucléico de PHDf_HD_IV de Populus tremuloides LG_XVIII1192

ATGGGTGATTCTGGAAAGAAATCAAAGCAGCAAGACTTGCACGAATCTTCACCATCTGACACAGTG
AATGGGTCGTTGCTGATTAAATCATTGAAGATAAAGAAGGGTGGCAAATTATCTCATAGAAAAAGC
GAAAAACCTAAAACTAAACCTCATCTGAAAACTATCATTAATTCATCTGTATCAAAGAAAAAGGTT
ACTCCTAAGAAGGGCATCAGGAATGGCTCTACCAGTAGAAGATTGATTCACAGGAAAATTCTGCAT
AAAGCACTTGATAAGAAGGCCTCAAGAAATGGGGCTTCCTCAGAGCTCCAAGGCAAACAGTTGTCA
ACTATTGATTCTGAGGGGAATGGAAAAAATGCCGATGAAGGTGCAATTAAAAAAGTCAAGAAGAGG
AAGCCTAAGAAAAGGCAAAAGGACAAGGTAAAGCTAGATGAACCACCACGCTTGCAGAGGAGAGCA
AGGTACTTGATGATTAAAATGAAGCTGGAGCAGAATCTTATAGATGCTTACTCTGGAGAAGGTTGG
AAAGGTAAAAGTCGAGAGAAGATTCGGCCAGAAAAGGAACTACTGAGAGCCAGGAAACAGATTTTG
AAATGTAAACTTGGATTACGTGAAATAATTCGGCAGGTGGATTCTCTAAGTACAGTAGGATGTATT
GAAGATGCTGTTATGGCTCCAGATGGATCTGTTTCCCATGAACATATATTCTGTGCGAAGTGCAAA
TTGAATGAAGTTTCCCAAGATAATGATATTGTACTCTGTGATGGGACATGCAACTGTGCCTTCCAC
CAAAAATGCCTTGATCCTCCTTTGGATACTGAAAATATACCTCCAGGAGATCAGGGATGGTTTTGC
AAGTTCTGTGACTGTAGGATGGAAATTATAGAAGCCATGAATGCTCATTTGGGGACCCACTTCTCA
GAGGACAGCGGTTGGCAGGATATTTTCAAAGAAGAAGCAGCAGTTCCAGATGGTGGAAATATGCTA
TTAAATCCAGAAGAAGAATGGCCTTCTGATGATTCTGAAGATGATGATTATGATCCAGAGAGGAGA
GAGAACGTCATGAGTGGAGCAGGTACTGATGATGATGCATCTGATGACACCAGCAATTCAACCCGC
TTGAGTTGGTCTTCAGATGGTGAGGTTTTTTCAGGATCCAGGAGGTGGGAGGTGGACGGCTTGGAC
TTCAGAAACAATTCTATTTATAGCAGTTTAGATTCTGATGAAACCAGTGATGGGGAAATTATTTGT
GGCCGTAGGCAGAGAAGAGCGGTTGATTATAAGAAGTTATACAATGAGATGTTTGGAAAGGACGCT
CCTGCACATGAGCAACCCAGCGAAGATGAAGATTGGGGTCCTAGCAAAAGAAAGCGACGAGAGAAG
GAGTCAGATGCAGCCAGCACCCTAATGACTCTATATGAAAGTAAAAGAAGATGTAAAAATGATGCA
ACCATTGAAGGCATGATGAAGCTTCCACGAGATCCTCAAATTCGAAGGCCAATTTTCAGGCTCCCC
CCTGATGCTGTTGAGAAACTTCGCCAAGTATTTGCAGAGAATGAACTTCCATCTAGAACTGTCAAG
GAGAATCTATCAAAAGAATTGGGCCTTGAACCTGGGAAGGTTAGCAAATGGTTCAAGAATTCCCGT
TATTTAGCTCTGAAGTCTAGAAAGGTAGAGAAGGGAGAACAAGTTCATTATTCTAGTTCCAAAGTC
TCTGCTGAACCCACTTTAAATGTCATGAAGGACAAAACTGCTGATCTTTCACTGTTAAAGGATAGC
CAGGCAGAAACTGGGGTGTGTACCCCAGAGAATTTGAAAAGGATCTTGCAGAGGAAGAAGCCGAGG
TCAATAAGTAAAAGTTTAAAGAAAAATGAACAGAAAAGAGGTTCCTTCGAATCACCTACTAAAAGC
AATGAGATGAATGTTGAGCACAATGATGATCTGAGCTTGAAGAAGCTGTTAAAAGCAAAAACAAAG
GGAGTAAAGAAAAAGGGTAATCGTATTTCAGCAGCTGCTGAATCTGATATGGAGAAACTTTGTAGA
GCAAAGACGAGAGTAGAGAATCTGAAGCAGAAACTGGTGAAACTACAAACTGGCAAAGCTAGGAAG
TCTTCCAAAATCCGTCCGCTAGACGAATCTGTTGTCTACGTTCCTATTGCTGAGCTAAGAGAAAAG
AAATGA

SEQ ID NO: 220, sequência de polipeptídeo traduzida de PHDf_HD_IV de Populus tremuloides

MGDSGKKSKQQDLHESSPSDTVNGSLLIKSLKIKKGGKLSHRKSEKPKTKPHLKTIINSSVSKKKV
TPKKGIRNGSTSRRLIHRKILHKALDKKASRNGASSELQGKQLSTIDSEGNGKNADEGAIKKVKKR

FIGURA 19 CONT.

KPKKRQKDKVKLDEPPRLQRRARYLMIKMKLEQNLIDAYSGEGWKGKSREKIRPEKELLRARKQIL
KCKLGLREIIRQVDSLSTVGCIEDAVMAPDGSVSHEHIFCAKCKLNEVSQDNDIVLCDGTCNCAFH
QKCLDPPLDTENIPPGDQGWFCKFCDCRMEIIEAMNAHLGTHFSEDSGWQDIFKEEAAVPDGGNML
LNPEEEWPSDDSEDDDYDPERRENVMSGAGTDDDASDDTSNSTRLSWSSDGEVFSGSRRWEVDGLD
FRNNSIYSSLDSDETSDGEIICGRRQRRAVDYKKLYNEMFGKDAPAHEQPSEDEDWGPSKRKRREK
ESDAASTLMTLYESKRRCKNDATIEGMMKLPRDPQIRRPIFRLPPDAVEKLRQVFAENELPSRTVK
ENLSKELGLEPGKVSKWFKNSRYLALKSRKVEKGEQVHYSSSKVSAEPTLNVMKDKTADLSLLKDS
QAETGVCTPENLKRILQRKKPRSISKSLKKNEQKRGSFESPTKSNEMNVEHNDDLSLKKLLKAKTK
GVKKKGNRISAAAESDMEKLCRAKTRVENLKQKLVKLQTGKARKSSKIRPLDESVVYVPIAELREK
K

SEQ ID NO: 221, sequência de ácido nucléico de PHDf_HD de Ostreococcus tauri CR954214.4

ATGAATGAGGTGCGCGACGCGGAGGAACGATGTAAACGTCGTGTGCGAGCGGTCGCGGACGCGGTG
CGGAAGGGAGGCGTGGTGTGGGCGATCGTGGGGTGGAAAGATAGTCAATGCGCGCGCGCGGCGTCC
GACGCGCGCGTTGTTGGAGATTGGGCGCGCGCGATGTGTGTGAGANGACTGACGATGCATCGATCG
GTGGTTGATGATACGNAGTTTGCGCTCGCGCGATCGCGGATACAGTCGCAGTTCGCGACGATTCGA
AGACATCAGGCGTTGGTGGAAGCGTACGCGAGCGACGGCTGGCGCGGACAGGCGGCTCAAAAGCCG
AAACCAGTCAGGGAGATTGAAAAGGCGCGCGAGAAGATATTCGAGGGTAAATTGAAGATACGGGAG
TATTTTAAGGTTTTGGAGTTCGATGAGCGCGAGCGCGAGATCACGACGGTGGCGGACGAGTTTGGG
GAGTGCGACGCCGCGGATATTTTTTGTAGCAAGTGCACCCTCGCAGACGACAGGCATGACGACGAT
ATTTTGCTCTGCGATGGCTTCTGCGATCGTGCGTATCATCAATCATGCGTGGCTCCACCGGTTTTG
GCGGAAGACATTCCGCCCGAGGACGAAGGTTGGCTCTGTCCGAGATGCGACGCGCGAGTGGACGTC
ATTTACGTCTTGAACGACGAGTACGATCAAAATTTGGGCCAGAGGTGTGTTTCAGCGGATATTTTC
GTTGCAGAGGCGGACATGCGAGATAAAGGGATCGTTCCAGGAACGGCGCAGTTCAAACACGCGCAC
GAAGAGGATTGGCCGAGCGATGAGAGCGATGACGAGGACTTCGATCAGGGCGGACACAGCGACGAT
GGGCGAGACGACGAGCACGAGGCGCTGAGCGGAAGCGCACAGTCATCGAGCGATGAATCGAGCAGC
GAATCAGAGTCTGATCTCATCATCGAGGGCCCACGAAGGCGCACGAAGGTTGATTACGTCGCCCTG
AACAACGCCATGTTCGGCGACAGCGAGGCGTACGAGGGCGAAGCTGAGGAGCTCGGTTGGAAACGC
AGCAACAAAACCAAGGCGGAGATGCTCGCGGCGCTCAAGGGTGAGAGCGCGAACGGGATAACGAAC
GGTGCCACGAAGAAGGATCCAAAATCAAAACGCTCGAAGCGAGAGCGCGTGGATGAGAAATCTCCC
GCACCGCGCGTCAGGTTCACTCCCGATCAAAAACTTGAGCTCGAGCGTGTCTTCTCCGAGTGCGCG
TCGCTCGACATTGACGCGCGCGACGCCCTCGCGAAGCGACTCGGGATACCCGGCGGCGCAAATTCG
ATTAAAATATGGTTCATGAATCGCCGTCGTAGGGCTAAGATTTTGGGTGAGAGGNAATACNGACAG
CACGATAATCCACGTCCCGCTCGCTCGTTCGACGCCTCGTCACAAACAACCTCTGAAAAACCTTCA
AGTCACCTCAGAAGCTCACGCGCGTCTCAAACATGTCGATCGCCCCTAAACGCGCCCCCAAACTCA
TATACTTCTCCGCGTGGTTCTGCCCGTTCGCCCATCGAGCGACGCTCGCGTTGGAGCATTACGCGC
ACGGCGGAGAGATTTCCTACGAGTGGAAGGAATCGCTCGGTGAGGTCACGCTTCGTTGACATTNGA
CGCTCGAATGCGACGAAACTTGAGCGCCTCGACNGACCGGAGACGAATCGCCGCGTAGGATGGGAA
AAACGCAAGTCAACCTCGAGTGAGGTGAAGACGAAACACGAGAATTGGTACCATTACAAGTCCCCA
GAGCTACTGAGACACAACCCGCTGGGCATGGTGCCGACACTCGTGACGCCGGCGACGTATGAGGAT
AAAGATGGAGATCCATCGCGGAACGTGGTGACGGAGAGCTTGGTGTGCGTACAGTTTATCGATGAA
CTCGTTCGGGGAAACGCGTACGCCGGGGAGACGGCGGCGATCATGCCGAGCGACCCGTACGCGCGC
GCGAAGGCGAGAGTGGACGCCGACTGGGTGAACAAAAACGTATGCTCGAGGTATTATCACGTCCTG
GTGCGACAAGAGGCGAGCGAGCAGAGAGAAGCGTTTGAAAAGCTCGTGGAGGGGCTGGAGACATTC
GCGGCGTGGTGCGGCAGCGGAGAAGGGAAGTTCTATGGTGGACAAACGACGCCCGGATTGGTGGAC
TACGCGCTCTTTCCATGGGCGTGGCGCCTCCCGGTGTTCGAGCACTACCGCGGGATGGATTTCAAG

FIGURA 19 CONT.

ATTCCTCGCACCGCGGCACTGAAATCGTATCACGATTGGATGGAAGCCATGCTCGCGCGAGAGCAA
GTTCGCAAGACGCTTCCACCGTGGGACGATTACTTGGAACACATTGGACGATACGCCGACGGGAGC
GCGCGATCGAAAGTCGCCAACGCAGTCCGCAGCGGCCGTGCGGCGCACGATTACGACGACAAGGAG
GACAACACCGATAGTTGA

SEQ ID NO: 222, sequência de polipeptídeo traduzida de PHDf_HD de Ostreococcus tauri

MNEVRDAEERCKRRVRAVADAVRKGGVVWAIVGWKDSQCARAASDARVVGDWARAMCVRXLTMHRS
VVDDTXFALARSRIQSQFATIRRHQALVEAYASDGWRGQAAQKPKPVREIEKAREKIFEGKLKIRE
YFKVLEFDERERETTVADEFGECDAADIFCSKCTLADDRHDDDILLCDGFCDRAYHQSCVAPPVL
AEDIPPEDEGWLCPRCDARVDVIYVLNDEYDQNLGQRCVSADIFVAEADMRDKGIVPGTAQFKHAH
EEDWPSDESDDEDFDQGGHSDDGRDDEHEALSGSAQSSSDESSSESESDLIIEGPRRRTKVDYVAL
NNAMFGDSEAYEGEAEELGWKRSNKTKAEMLAALKGESANGITNGATKKDPKSKRSKRERVDEKSP
APRVRFTPDQKLELERVFSECASLDIDARDALAKRLGIPGGANSIKIWFMNRRRRAKILGERXYXQ
HDNPRPARSFDASSQTTSEKPSSHLRSSRASQTCRSPLNAPPNSYTSPRGSARSPIERRSRWSITR
TAERFPTSGRNRSVRSRFVDIXRSNATKLERLDXPETNRRVGWEKRKSTSSEVKTKHENWYHYKSP
ELLRHNPLGMVPTLVTPATYEDKDGDPSRNVVTESLVCVQFIDELVRGNAYAGETAAIMPSDPYAR
AKARVDADWVNKNVCSRYYHVLVRQEASEQREAFEKLVEGLETFAAWCGSGEGKFYGGQTTPGLVD
YALFPWAWRLPVFEHYRGMDFKIPRTAALKSYHDWMEAMLAREQVRKTLPPWDDYLEHIGRYADGS
ARSKVANAVRSGRAAHDYDDKEDNTDS

sequência de ácido nucléico parcial Aquilegia formosa x SEQ ID N°: 223, Aquilegia pubescens de PHDf_HD contíguo de DR914726.1, DR941696.1, DR943570.1

ATGCAAGATGGTGGGAACCTTGGTTTTCCAAAGAAACGTGTTGGATCCCAAATAAATACCTCATTT
CAGAAGAAGAATAGTTGCAAAAATCCGCATATTAGGAGATGGAAACCAAAGCCAAGATCTCATGTC
AAGACAATTGCTTCGTTCTTGTCGAAGAGAAAGACATCTGAGCTCTCAAGCAACAGAGCCAGAAAT
AAATGTTCCAATGGAAAACCTTTTAGAAGAAGAGCGTCACAGAAACCCAATGATACCAGTCCTTCT
GATAGGCCACTTTCATCAATGCTTAAAGATTTCAAATCGTTAAGCAATAAATGTCAGGGACGTGAA
AAGACAACTGACAAGGATCTTGGAAGTCAAAAGGCTCACAGAAGGAGAAGGAGCAAGAAGAGGAAG
AATGAGGAGTTAGATGAAGCCGCTCGAGTCCAAAGAAGAACAAGGTACCTCCTGATTAAAATGAAG
CTGGAGCAGAACCTTATTGAGGCTTACTCTGGGGAGGGTTGGAAAGGCCAGAGTCGTGAGAAAATT
AAGCCAGAAAAGGAACTGCAAAGAGCCAAGAAGCAGATATTGAAGTGCAAACTTGGAATTCGTGAT
GCTATTCGTCAGCTGGAATCTCTTAGTTCTGAGGGAAGCATTGAAGATTCTGTGATAGGTTCAGAT
GGATCTGTTTTCCATGAACATATTTTCTGTTNNNNNNNNNNNNNNNNNNNNNAGCATTTCCAGAT
AATGACATTATACTGTGTGATGGAGCATGCAATTGTGCTTTCCACCAGAAATGCCTGGAGCCTCCA
TTAGCCACTGAAAATATTCCTCCTGGTGATCAAGGGTGGTATTGCAAATTTTGTGAGTGCAAGATG
GCAATACTTGATGCAATCAATGCACATCTTGGGACTCAGTTTTCTCTAGACAGTAAATGGCAAGAT
ATTTTCATGGAAGCAGCTAATGCACGCGATAATGAGGATATCTCAATTCATCCTGATGGTGACTGG
CCATCTGATGACTCTGAAGATGTTGATTATGACCCAGAAAAATATGAAATTGATTGCAGCTATAGC
AGAATGGGCTCGGAGCACAACGTGTCTGACGATGCAAGTAGTTCCGGCACCTTGTATTGGACTTCC
GAAGAAGATGATGGTTTGCATTCTGAGCGATTCAGAACTGATGGTAGTCTGGAGGGGGGATTTTAT
GGGGATAAAACTAAAATCCAGAGTTCTGATAGCATAGATTCCAATGACAGTACTGATGGTATGGCA
TTTCATCGTAGGCAGCGAAGAGATGTTGATTATAAGAAACTGCATGATGAAATGTTTGGTAAAGAT
ATGCCTGAAAGTGAAGAAATCAGTGAAGATGAAGATTGGGGACCTAGAAGAAGAAAGCGCAGAAGA
AATGAGTCTAATGCAGCCACTACTTTAGTATCTTCGTGTGGTACTGAGAATGAATTTTCGAATGTG
GCACTTNNNNTACCATCAAGAGTGGTCAGGGAGAGTCTCTCGAAGCAATTGGGTATTGCAATTGAG
AAGATCAACAAGTGGTTCAAGAATGCTCGGTATACAGCTTTAAAAATTAGAAAGGCAGAAAAGACT

FIGURA 19 CONT.

GAACAATCCCCTGGAGGTGGCAAGATGCAAAGTGCTGACGTATTAACATCAAAGGAAAACCATTAT
TTGGTCTCAACAGGAACAGCAGTCCAAAGATCCAAAGTTTCGAGTGTCCATAGAAGTAAAAACCGT
AAATTGACAACCACACCCTCAAGGGAAGTACAAGAGAAAGTGTCTGCTGTTGCACCATCCGCTACT
GCAAATGAGGTGAGTGTACAGTTAAGAAAAGCAGTGAGCGTAGAGAAGGAAATGGGCAGTGTAAAA
AGAAAATCTGCACAAGTAAAGAGGATCAAGCGCACATGCTCCACTAAGGAGCAGCAGCTATACATG
ATTGAGTTGGAGAGAATTTGCCGTCTTGAGGAGAAGCTTGAAAAGATGAAGAAGGTATTGCTTAGT
ATAGATGACAATCATAGTAATGTATCAGATGGACCCCATATGAACGAACAGTCTGCCGTATACATT
CCAGTTGCAGAGCTAAGAGAAAAGATGTGA

SEQ ID NO: 224, sequência de polipeptídeo traduzida parcial de PHD_HD de Aquilegia formosa x Aquilegia pubescens

MQDGGNLGFPKKRVGSQINTSFQKKNSCKNPHIRRWKPKPRSHVKTIASFLSKRKTSELSSNRARN
KCSNGKPFRRRASQKPNDTSPSDRPLSSMLKDFKSLSNKCQGREKTTDKDLGSQKAHRRRRSKKRK
NEELDEAARVQRRTRYLLIKMKLEQNLIEAYSGEGWKGQSREKIKPEKELQRAKKQILKCKLGIRD
AIRQLESLSSEGSIEDSVIGSDGSVFHEHIFCXXXXXXXXAFPDNDIILCDGACNCAFHQKCLEPP
LATENIPPGDQGWYCKFCECKMAILDAINAHLGTQFSLDSKWQDIFMEAANARDNEDISIHPDGDW
PSDDSEDVDYDPEKYEIDCSYSRMGSEHNVSDDASSSGTLYWTSEEDDGLHSERFRTDGSLEGGFY
GDKTKIQSSDSIDSNDSTDGMAFHRRQRRDVDYKKLHDEMFGKDMPESEEISEDEDWGPRRRKRRR
NESNAATTLVSSCGTENEFSNVALXXPSRVVRESLSKQLGIAIEKINKWFKNARYTALKIRKAEKT
EQSPGGGKMQSADVLTSKENHYLVSTGTAVQRSKVSSVHRSKNRKLTTTPSREVQEKVSAVAPSAT
ANEVSVQLRKAVSVEKEMGSVKRKSAQVKRIKRTCSTKEQQLYMIELERICRLEEKLEKMKKVLLS
IDDNHSNVSDGPHMNEQSAVYIPVAELREKM

SEQ ID NO: 225, sequência de ácido nucléico parcial de PHD_HD de Glycine max Contig GM06LC25006

AAGAAGAAAAGACAAAGGAATAACATAGATGTTGATGATGCTTCACGCTTGCGAAGGAGAACAAGG
TACCTCTTAATCAAAATGAAGCTAGAGCAGAACCTTATTGATGCTTACTCTGGAGAAGGTTGGAAA
GGTCAAAGTCGGGAAAAGATTAGGCCAGAAAAGGAGCTACTACGAGCAAAAAAGCAGATTTTAAAA
TGTAAACTTAGTATTCGCGATGCCATACACCAGCTGGATTCTCTTAGTTCTGTGGGTAGCATNGAA
GATTCTGCCATTGCTCCAGATGGATCTGTTTATCATGAAAATATATTCTGTGCCAATTGCAAACTA
CATGAAGCTTTCCCAGATAATGATATTATACTCTGTGATGGCACTTGCAATCGTGCTTTTCACCAA
AGATGTCTCAATCCTCCTTTGGATACTGAAAATATTCCTCCTGGAGATCAAGGCTGGTTTTGCAAG
TTTTGTGAATGTAAGATAGAAATACTAGAGGCAACAAATGCCCATCTTGGGACTCAATTCTCCCTA
GACAGTAATTGGCAGGATGTATTCAAGGAAGAGGCTGCTATGCCTGATGGTGACATTGCATTACTA
AATCCTGAAGAAGAATGGCCATCAGACGATCCTGAAGATGATGATTACAATCCAGAGAGGAAAGAA
GACAACCACAACTTTGACACAGAAGGAGCTGATGAAAATGCATCCAATGATTCAACCAGTTGTTCC
AGTCTGTTGTCTTTGAATGGCGAATGTCCTCCAGTAGATGAAGGCATCTGTCATGAATATTATTCT
GTTAATAGCTGCTTAGATTCTGATGAATCTGGGGAAATAGCATGTGGTCCTAGGCAGCGTAAAGCT
GTTGACTACAAGAAACTCTATGATGAAATGTATGGGAAGGATGCTCCACCTTGTGAACAAATGAGT
GAAGATGAAGACTGGGGTCCTGGCAAAAGAAAGCGAAGAGAAAAGGAGTCTGATGCTGTTGATTCT
CTTATGACTTTGCATGAAAGTGAGAATAGGCATCCCAATAATGAGCACAACATGACAAGTAAAGAT
TCTTCAAGCATAAAGATCAAAAGGCATTGTTTTAGGATTCCACGTGATGCAGTTGAGAGGCTTCGC
CAAGTTTTTGCTGAGAACGAACTTCCTCCAAGATCTATAAGAGAGGGTCTTTCAAAAGAGCTGGGC
CTTGACACCGAGAAGGTCAGCAAATGGTTCAAAAATGCACGTTACTTAGCACTTAAAAACAGAAAG
CATCAAGCAGAAGGAGGAGCAGATCAGCTTCAGAGTATCACTTCTACAAAGAACAGATTACAAAAA
CAGGAGAATGTCGATCCTTTGAAATCAAAGACCCCAAAAATTACTAGGACACATTCCCAGAAGGAT
GTTAAAAATGTCAATGGAAGAAAGAAAATGAAGTCATCTAACAAGAAAAGACAACCAGAAATTCCT
CCACCTCCAGGAGAAAATGGCAAGAAGGACTTTATGGAGATCAGTGATGATGTGAGCTTGAAGAAA

FIGURA 19 CONT.

CTGTTGAAGAAAAGAAAGAAGAGGTTGATAAATTTTACGTTTGGAGGAGACTCTCAGTTGGCAGAG
TTGGAATTTGAAAGACTAAGCGAATTGAAGACTAAAGTAGATAGTATGAAGCAAAAATTAACTGCA
ATTCAAAATTACAGAGTTAAGGGTTCACCCTATTCGAATGAACCATCCATTGTATATGTACCCACA
GCTGTGTTAAGGGAAAAGGTTGAATGA

SEQ ID NO: 226, sequência de polipeptídeo traduzida parcial de PHD_HD de Glycine max

KKKRQRNNIDVDDASRLRRRTRYLLIKMKLEQNLIDAYSGEGWKGQSREKIRPEKELLRAKKQILK
CKLSIRDAIHQLDSLSSVGSXEDSAIAPDGSVYHENIFCANCKLHEAFPDNDIILCDGTCNRAFHQ
RCLNPPLDTENIPPGDQGWFCKFCECKIEILEATNAHLGTQFSLDSNWQDVFKEEAAMPDGDIALL
NPEEEWPSDDPEDDDYNPERKEDNHNFDTEGADENASNDSTSCSSLLSLNGECPPVDEGICHEYYS
VNSCLDSDESGEIACGPRQRKAVDYKKLYDEMYGKDAPPCEQMSEDEDWGPGKRKRREKESDAVDS
LMTLHESENRHPNNEHNMTSKDSSSIKIKRHCFRIPRDAVERLRQVFAENELPPRSIREGLSKELG
LDTEKVSKWFKNARYLALKNRKHQAEGGADQLQSITSTKNRLQKQENVDPLKSKTPKITRTHSQKD
VKNVNGRKKMKSSNKKRQPEIPPPPGENGKKDFMEISDDVSLKKLLKKRKKRLINFTFGGDSQLAE
LEFERLSELKTKVDSMKQKLTAIQNYRVKGSPYSNEPSIVYVPTAVLREKVE

SEQ ID NO: 227, sequência de ácido nucléico parcial de PHD_HD de Lotus corniculatus AP004517

ATGCAGGGTACTGAGAAGTTAAATGGCACAGGATCTACAAAATCCAGTAACTCAGAGGAACATGCT
GAGTCAAAGGTAGATTCATTGAGATACAGAACAAATATAACAAAATGCCGAGGGAAGAAACAAAAA
CTGAAATCAAAATCTCATAAACTCAGCGGTTGCTTAAGTGCATCAGGGAGGACAGTTACCAATTCT
TCCATCAAGAGACAAGTAAAGGATTCTTCTAATAAAAAGATTATTAGACAAAGTCTACATAAAACT
GATGGCAAATCTTCACGGATGCTGTCTTCAACAATGATACAAGGAGGAAAATCTTCTCTTGGTTCT
AGAAAGGAGGGCAAAGATGTTGATGGGGAGGTAATGGTTCAAAAAGTTAAAAGAAAGAGAAAGAAG
AAAAGGAGACAGGATGATGTTGATCTTGATGATGCTTCACGCTTACTGAGAAGAACAAGGTACCTC
TTAATCAAAATGAAGCTAGAGCAGAATCTTATTGATGCTTACTCTGGAGAAGGTTGGAAAGGTCAA
AGTCGAGAGAAGATTAGGCCAGAAAAGGAGCTACAAAGAGCCAAGAAGCAGATTTTGAATTGTAAA
CTTAGTATTCGGGATGCCATTCGCCAGTTAGATTCTCTTAGCTCATTGGGCAGCATTGAAGATTCT
GTCATTGATCCAGATGGATCTGTTTATCATGAACATATATTCTGTGCAAATTGCAAGTTGCGTGAA
GCTTTCCCAGATAATGATATAATACTCTGTGATGGCACATGCAATCGTGCTTTTCACCAACGGTGT
CTCAACCCTCCTTTGGATACTGAAAATATTCCTCCTGGAGACCAAGGCTGGTTCTGCAATTTTTGT
GAATGTAAGATAGAAATATTGGAGGCAACAAATGCCCATCTTGGGACTCAAATCTCCTTGGACAGT
ACTTGGCAGGATGTATTCAAGGAGGAGGCTGCTATACCCGATGGTGATAATGCATTACTGAATCCA
GAAGAAGAATGGCCATCAGATGATCCTGAAGACGATGATTATTATCCAGAGAGGAAAGAAGACAAC
CACAGCATCAAAGCAGAAGGAACTGATGATAATGCATCCAATGATTTAAGCAGTTCTTCCAGTCTA
GGGTCTTTGAATGGTGAATGTTCTCCAGTAGATGAAGGCACGGGTCTTGAATATTATTCTGTTAAT
TGCTGCATAGATTCTGATGATTCTGGGGAAATAGCATGTGGCCGTAGGCAGCGTAAATCTGTTGAC
TACAAGAAACTATATGATGAAATGTTTGGGAAGGATGCTCCTCCATGTGAACTAGTGAGTGAAGAT
GAAGACTGGGGTACTGGCAAAAGAAAGCGAAGAGAAAAAGAGTCTGATGCTGTTAATACTCTCATG
ACTATGCATGAAAGTGAGAACAAAAATCCCAATAATGAGAATAGTGATGGAATAAGAGAGGGTTCT
TCAGGCAAACCAATAAGAAGGTCTTGTTTCAGGATTCCACATGATGCACTTGAGAAGCTTCGCCAA
GTTTTTGCGGAGAATGAGCTTCCTCCAAGATCTGTCAAAGAAGGTCTTTCGAAAGAGTTGGGAATT
GATGCTGCGAAGGTTATGCCTTACAACCTC

FIGURA 19 CONT.

SEQ ID NO: 228, sequência de polipeptídeo traduzida parcial de PHD_HD de Lotus corniculatus

MQGTEKLNGTGSTKSSNSEEHAESKVDSLRYRTNITKCRGKKQKLKSKSHKLSGCLSASGRTVTNS
SIKRQVKDSSNKKIIRQSLHKTDGKSSRMLSSTMIQGGKSSLGSRKEGKDVDGEVMVQKVKRKRKK
KRRQDDVDLDDASRLLRRTRYLLIKMKLEQNLIDAYSGEGWKGQSREKIRPEKELQRAKKQILNCK
LSIRDAIRQLDSLSSLGSIEDSVIDPDGSVYHEHIFCANCKLREAFPDNDIILCDGTCNRAFHQRC
LNPPLDTENIPPGDQGWFCNFCECKIEILEATNAHLGTQISLDSTWQDVFKEEAAIPDGDNALLNP
EEEWPSDDPEDDDYYPERKEDNHSIKAEGTDDNASNDLSSSSSLGSLNGECSPVDEGTGLEYYSVN
CCIDSDDSGEIACGRRQRKSVDYKKLYDEMFGKDAPPCELVSEDEDWGTGKRKRREKESDAVNTLM
TMHESENKNPNNENSDGIREGSSGKPIRRSCFRIPHDALEKLRQVFAENELPPRSVKEGLSKELGI
DAAKVMPYNL

SEQ ID NO: 229, sequência de ácido nucléico parcial de PHD_HD de Sorghum propinquum BF656332

AACTCGAGAACAAATGAAGAAAAATCTGGTATATCAGGGAAGATTGATTCAGTAGACAACTCTTGC
TTGGTTCCCTTGTCTGAAATCATCAATGTGCATACACGAGTGCACCACAATCTTGAGATGAGGAAG
ATGGAATCAACTAGCAGTCCTGTGTGGCTGCACAATGAAGGAGGCTGCTTGTTTCCAACACTCCAA
GCCAAGGAGAGTACATTACCTACAAGCAAGCCTTGTTTGCCGTCAGAAATAACCCATCCAACAACT
AATGAAGTGAGCACTTTAGTGCAAGCTACTTCTTGGATGGATGCTGGGGCGTGCATTGAGCAGCAA
GAAACCACTCCTTGGGTGGATACCGGGGCCTCAGGTTACCAGCCTTTCCTGGACGTGATTGATGAG
ATGTGCGGACTTGAGTGCAGGCTGCAGAGGCTGAAGGAGAACATGTTCTCATCTGGCATGGACGGC
AGAACCGCCGGTGTGAGTGACATGGGAAACCAGGCTGTGGTGCTTGTGCCGACTGCTGAGCTCAAG
GAAAAAGCGCCGCATGGCGGTTTATTTGGGCACTATTGCCCATAG

SEQ ID NO: 230, sequência de polipeptídeo traduzida parcial de PHD_HD de Sorghum propinquum

NSRTNEEKSGISGKIDSVDNSCLVPLSEIINVHTRVHHNLEMRKMESTSSPVWLHNEGGCLFPTLQ
AKESTLPTSKPCLPSEITHPTTNEVSTLVQATSWMDAGACIEQQETTPWVDTGASGYQPFLDVIDE
MCGLECRLQRLKENMFSSGMDGRTAGVSDMGNQAVVLVPTAELKEKAPHGGLFGHYCP

SEQ ID NO: 231, sequência de ácido nucléico parcial de PHD_HD de Sorghum propinquum BF704605

AAGATAAAATTGGAGCAGAATCTACTAGATGCTTACTCTGGAGATGGATGGAATGGGCAAAGCCGG
GAGAAAATAAAGCCAGAGAAGGAACTGCAGCGTGCCAGGAAACAAATCATAAAATGCAAAATTGCT
ATACGTGATATCATTCGCCAACTTTGTTTGTATACTTCTACTGGGAGTGTTGATGACCCAGCAATG
CCTCCAGATCAGTTTACCAATCCCGAACATACCATGTGCTCAACATGCAAGTCTCATGAATCATTT
CCCAGCAATAAATTTATCTTCTGTGAAGGGCCCTGCAAAAGGGCATATCATGAGAAATGTTTGGAA
CCTCCCTTAAACAAAGGTGTACTTCCAACAAGTAGCCATGGATGGCTTTGCAAATTCTGTTTGTGC
AAGGTGAAAATTTTAGAAACTATTAATGCACATCTAGGAACAAGTTTTACAGTGATGTGCTCCTTT
GAAGATATATTCAAGGAAGCAACAGAACAAATAGACTCCGAGGATGCACTAGATGAAGATTGGCTC
TCTG

SEQ ID NO: 232, sequência de polipeptídeo traduzida parcial de PHD_HD de Sorghum propinquum

KIKLEQNLLDAYSGDGWNGQSREKIKPEKELQRARKQIIKCKIAIRDIIRQLCLYTSTGSVDDPAM
PPDQFTNPEHTMCSTCKSHESFPSNKFIFCEGPCKRAYHEKCLEPPLNKGVLPTSSHGWLCKFCLC
KVKILETINAHLGTSFTVMCSFEDIFKEATEQIDSEDALDEDWLS

FIGURA 19 CONT.

SEQ ID NO: 233, domínio de dedo de homeodomínio de zíper/planta (ZIP/PHDf), compreendido na SEQ ID N°: 180

DEVSRMEKRARYLLIKIKQEQNLLDAYSGDGWNGHSREKIKPEKELQRAKKQIMKYKIAIRDVIHQ
LDLCSSSGSKDDSVIPPDGCHESVNPEHTICSRCKSHESFPDNNIIFCEGGCKLACHQKCLEPPFD
KILPTTRHGRLCKHCSSKMKILDAINAHLGTSFTVKCPSSDIFKEAAE

SEQ ID NO: 234, homeodomínio, compreendido em SEQ ID N°: 180

RIPPAAVEVLRKAFAENELPARSVKENLSTELGISFEKIDKWFKNTRCAA

SEQ ID NO: 235, promotor GOS2 de Oryza sativa

AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAATAT
AAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCACCTACT
TTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGTAATTAAGT
GGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGTTAAATTATTCG
AGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTTCTAGCTGAACTCAATGGGT
AAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTCTGAACGTATTGGCAAAGATTTAAA
CATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACATCATTAAGGACATGTCTTA
CTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATTTTTATTATTTATCTTTTTTCG
ATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAAT
ACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTAGGTAGCAATATC
TGAATTCAAGCACTCCACCATCACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTT
ACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAAAAAAGAATT
TTGCTCGTGCGCGAGCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCCACA
GAACAACCCACAAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAG
GCTTTGCGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGCATCCTCCTTCTCCCATCTATAA
ATTCCTCCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAAG
GACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTCGATCCATATCTTCCGGTCGAGTTCTTGGT
CGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCTCCCTTCGGTTGTTCTTGGATTT
ATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCC
TGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTCGGTTTGATTAGTAGT
ATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTAGGGATCGGAATCTTGCGATT
TTGTGAGTACCTTTTGTTTGAGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAGTAC
GGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTC
CCTATTGAACAAAAATAATCCAACTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTT
AAGCCTGTCCAAAATTTCGCAGCTGGCTTGTTTAGATACAGTAGTCCCCATCACGAAATTCATGGA
AACAGTTATAATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTT
TTTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATG
CTTTTATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGA
AGAACTTATCCGATTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATT
CATTTGGATTATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAA
CTGTCCTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGT
AGAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAATCG
GGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCTTGCCACT
TTCACCAGCAAAGTTC

SEQ ID NO: 236, prm09687

GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGAATACCCCAGAAAAGAAA

FIGURA 19 CONT.

SEQ ID NO: 237, prm09688
GGGGACCACTTTGTACAAGAAAGCTGGGTGATGCAAGGTTAAGATGCTTT

SEQ ID NO: 238, sequência de ácido nucléico de Phypa_PHD-HD XM de Physcomitrella patens XM_001762483.1
ATGATGAAGAATTACCGTCCTCCATCACCACTTCCAAGTGCTCGCCAAGCTTTCTTACAAGACCCG
GCGTCGTCATTCCAAATGACAGGACGACATTGGGCAGACCCAGAGGGAGGTGGAGAACGGAAGCTT
CACACACACTCTCACTCACACTCACTCACTCACGAAGGAAACGAAGGGCAAAGGAAGGGGAGCCCC
GCCGTGAGCCGAGAGCGGAGAGCAAGCATTCCCAACGGAAGTCTCCGTCAATGCGGTGCCGGAGCC
GACGACGACGACGATGATGGGCTGAGCCTGTCATCGCAATGGATGCATTGCCCTGTAAGTGGGAAG
GTGGAGGAAGGATTGGGGAAGAGGTTGAGGAAGAGAAAGAGAAAGAGAAAGAGTTTGTTGTTGTTG
TTGGGGAAGGGGAGGGGGCGACGGATGATGCCCATGGGCGTAGTCCGCGACGATGAGTCCCAGGCC
GCTGCTGGCTCGCCGCCTGCGCCCACCCCTCCCCCAATCGCCAAGCCGACGCTGCGATCCCACTCG
CGCTCCAGGGCGCACAAGGCGCGCTCCCAACACCCGCCGCTGGGCGATGGCGATGCGCCCTTGCCC
CCACCGCAAGGAACAGCAACACCAGGAGGAGGAGGAGCAGCAGCAGTGGGCGCACCAGGGGCCATG
GATGCCGAGGCTATGGCGGCGGGGCAGCGGCGGCGTCGACGACGCCGACGACAGCGGGGCAAGAAC
TCGGATGTGCCGCTTGACAATCGCACGCGAATTAAGAGGCGTGTCAAGTACTTGCTCATGAAAATG
CGAGTCGACCAGAACCTGCTGGATGCGTACTCCGGGGAGGGTTGGAAGGGGCAGAGGAATTCTCTT
TTTGGCATATCCAGTCGGGAAAAGATCAGGCCGGAGCAAGAACTACAGCGTGCTGAAGCGCAGATC
CTGCATTCGAAATTGGCCATACGGGAGGCTATTCATGAATTGGATAACTTGGGGTTAGAAGGCTCG
CTCGACGAAAATGCATTTGATGCTGAAGGCCGTGTTTACCATGAAGAGATATTTTGTGCCAAGTGC
CGATCTCAAGATGCTCTACCTGACAACGACATTATTTTATGTGACGGCGCATGCAACAGAGGCTTC
CATCAGTACTGTTTAGACCCACCTTTAGCTACGATCAACATTCCTCCTGGAGACGAAGGGTGGTTG
TGTCCTGTATGTGAGTGCAAGATGGAGTGCATTGAAGTTATTAATGCATACTTGGGCACACACTTT
GAGGTGGAGAATAGTTGGGAGGAGTTGTTTTCTGAAGAAGCTCTCGTGGCTGCGGGCGGTAGAACT
CAAGGGGTCACAGATGGTGATTTTCCTTCCTCAGACTCAGAAGACGACGATTATAACCCTGAAGGA
GCGGAGAAGCGGACAGACAGTGAAAGCGAAGGCGAAAGTGATGAAGGTGGGGGAGATAGTGATTCA
GGAAGCGATTTTAGGGAAAATGAAGCGTCTGGTAGTGACAAAGATGAGCAAGATCCAATGCAGGCA
ACTAATTTAGAGAGACTTGCGAGAACAGCAAAGCGGACCAGCAAGGAAGGAAGTGGAGGTCCAGAA
TCTTCAGGCAGTGATGACTCAGAAGTTTTTCTTGTTGATGGGTTTACGGATGGTGGAGTCAAATTA
GGTTTAAGACGTAGAGACTCACAAGCAGAAGAATCAGCTTCAGACGAGGAAACCATGGTTATAGCA
GGAAAAAGACATCGAAAAGCAGTAGACTACAAGAGATTACATGATGAAATGTTTGGGAACATGGAG
ATTGATGTGGACGATATCATATCCGAAGATGAGGATTGGGGTCCGAAACGCCGTCGGAGAAGAACC
ATTCCTGATGACCCAAGCAGGTCTAGACCTCCTCGAGTTCGACGGAGAAAGGCGCGAACCTCAACT
GGAGATGCTTTAAAAGAAGTTGATGCAGATGGGGTACCATCAGATTTACCCTCTGGGTCACAAGGC
GAGGGAGCAGCCGATACTCTTGAAGCTGTTAGAGACAAACGAATGTGGAGGAAGTTGCCTGATTCT
GCTGTTGAGACGTTACGACTGGCTGCGGCAGTGACGACACTGCCCTCAAAATCCCGTAAAGAGGAG
TTGGCGACACAGCTTGGCTTGAGTTTTTCACAAGTCCATGGATGGTTTAAAAATCAGCGGTATTTA
GCTCTCAGAAATGGGTTGGCGATTTCAAAGCGACCAGCAGCTGGGCGGATGGCGGCCACTGTAACC
ATACAATCGGAACACGACAACAATTCAAGTGCCTATTTAACTGAGAAGTTGGATGAAGTCCAAATG
CGACTTTTTGAACTGAAGCAAACTTTGGAAGATTTCGCTAAGAATTCATCTGGTTTAGGCGGTGGA
GAATCAGATAACAGTTTTGAAAATGGAGGAATTGGACTTCCTGCAAATGGGAAACGGATTATCTAT
GTCCCAGTGGCTGAGGTTAGAGAACGGCCTTCTCTTTTCTCTGACAGCTGTCCGCTTTCATTATTA
GCAGTGTTAAATTGCGAGGTGAATTACAAGAGGAATCTACGTATATTCATGTGTTCATGGCCTAGC
TGCTTAGCGTACGGGGTTTTGCATGCCAACTGCCCTTGTCATATTGGATCATGTTTCGAAGGAATA
GGTGGGTGA

FIGURA 19 CONT.

SEQ ID NO: 239, sequência de polipeptídeo traduzida de Phypa_PHD-HD de Physcomitrella patens

MMKNYRPPSPLPSARQAFLQDPASSFQMTGRHWADPEGGGERKLHTHSHSHSLTHEGNEGQRKGSP
AVSRERRASIPNGSLRQCGAGADDDDDDGLSLSSQWMHCPVSGKVEEGLGKRLRKRKRKRKSLLLL
LGKGRGRRMMPMGVVRDDESQAAAGSPPAPTPPPIAKPTLRSHSRSRAHKARSQHPPLGDGDAPLP
PPQGTATPGGGGAAAVGAPGAMDAEAMAAGQRRRRRRRRQRGKNSDVPLDNRTRIKRRVKYLLMKM
RVDQNLLDAYSGEGWKGQRNSLFGISSREKIRPEQELQRAEAQILHSKLAIREAIHELDNLGLEGS
LDENAFDAEGRVYHEEIFCAKCRSQDALPDNDIILCDGACNRGFHQYCLDPPLATINIPPGDEGWL
CPVCECKMECIEVINAYLGTHFEVENSWEELFSEEALVAAGGRTQGVTDGDFPSSDSEDDDYNPEG
AEKRTDSESEGESDEGGGDSDSGSDFRENEASGSDKDEQDPMQATNLERLARTAKRTSKEGSGGPE
SSGSDDSEVFLVDGFTDGGVKLGLRRRDSQAEESASDEETMVIAGKRHRKAVDYKRLHDEMFGNME
IDVDDIISEDEDWGPKRRRRRTIPDDPSRSRPPRVRRRKARTSTGDALKEVDADGVPSDLPSGSQG
EGAADTLEAVRDKRMWRKLPDSAVETLRLAAAVTTLPSKSRKEELATQLGLSFSQVHGWFKNQRYL
ALRNGLAISKRPAAGRMAATVTIQSEHDNNSSAYLTEKLDEVQMRLFELKQTLEDFAKNSSGLGGG
ESDNSFENGGIGLPANGKRIIYVPVAEVRERPSLFSDSCPLSLLAVLNCEVNYKRNLRIFMCSWPS
CLAYGVLHANCPCHIGSCFEGIGG

SEQ ID NO: 240, sequência de polipeptídeo traduzida de Phypa_PHD-HD II de Physcomitrella patens XM_001779822.1

ATGGTGCCGACGGGCGTGGTGGAGCGGGAGGACCAGCCTGAGCCTGCTGCTGCCCCCCCAGTGCCG
GTGATTACACCCAAGACGCTTCGATCGCACTCTATCAGGACGCAGACCACGCCAGGTGATGGCGAT
GCTGGTGCGGATCATGCTCCTGGAGCAGTGGCAGCAGCAGCAGCATCAGCAGGAGGAGGAGGAGCA
GCGGATGTCATGCAAGGTAACGATGCGGGGCAGGAACGGCGAGGACGACGGCGCGGGCGACGGAAG
AAGTCAGACGTGCCGCTCGATAATCCCACGCGAATCCTAAAGCATGTCAAGTACTTGCTCATTAAA
ATGCGAGTCCACCAGAACCTGCTCGATGCGTACACAGGGGAGGGTTGGAAGGGTCAGAGTCGCGAG
AAGATCAGGCCGGAGCAAGAACTACAACGTGCGAAGGCGCAGATACTGCGATCTAAGTTGGCCATA
CGGGAGGCTATTCATGAACTCGATGACGTGGGTTTAGAAGGTTCTCTCGACAAAAACGCTTTTGAT
TCTGAAGGTCGTATTTACCATGAAGAGATTTTTTGTGCGAAGTGCAAATCCCAAGAAGCTCTTCCT
GACAACGACATTATCTTGTGCGACGGGGCGTGCAACAGAGGCTTCCATCAGTATTGCTTAGACCCA
CCTTTAGCTACGGAAGACATTCCTCCTGGAGATGAAGGATGGTTATGCCCTGTATGCGACTGCAAG
ATGGAGTGTATTGAAGCGATAAATTCATATTTTGGCACCAGTTTTGAGGTCGAGAATAGTTGGGAG
AGCTTTTTTTCGAACGAAGCTGGTATAGCTGCAGGCGGTGGAACGCAAGAAGGGGCTGGTGGTGAT
TGGCCTTCCTCAGGCGATGAAGATGATGAGTATGATCCAGAAACAGCGGAAGGGCCTTCTTCAGCT
TCTGAAAGTGGAGCTCAAGTAATAGGGCTTGCTGTTGACTGGCCGTCCTCAGACGACGAAGATCAC
GACTATGATCCGGAAGCAGCGGAAGAACCTCCTGGAGCTTCTGGGAGTGGAACTCAAACAGTAGGT
CTTGACGTTGATTGGCCTTCCACAGACGATGAAGATGACGACTTTGATCCAGAAGGTATCAGAAAG
CAGGCAAACAGTGACGAAGAAAGTAGTAGTGATGAAAGTGAGGAGGACGGGGAGGATAGTGGTTCT
GTAAGCGGCTTCAGGGAAAATGAGGCGTCTTCAGGTGTAAGTGACAAAGATGAGCAGGAGCCTAGT
GAAAAACATAATCTGGAGCTAACTGCTGTCACTAACATCAAGAAACTTGCAAGAAAAGGAAAACGT
AACATAATCAGCAACAAAGGAAAAAGTGAATATTTAGAGTCTTCAGACAGTGATGATTCTGATGTC
TCTCTTTTTGACGAGTTCAAGGGTGGTAATGCCAATTTAGGCTTAAGACGTGGGGATCCACAGGCA
GCAGTATCAGCAGCAGACGAGGAAGCCGTGATTATTGCTGGGAAAAGACATCGGAAAGCTGTGGAC
TACAAAAGACTACATGATGAAATGTATGGTAAAGCGGAGGCCGATAACGATGACGACGTTATCTCT
GAGGATGAAGATTGGGGTCCTGAACGTCGCCGAAGACGAACCATTCCTGATGATCCTAACTCGCCA
AGGTCCCGAATTCGAGGTAGGAAAGTGAGAACTCCAAGTGGAGATATTTCTAAAGGAGCCGATGCA
GATTTACCCTCTGGATTACCAGATATGCCATCCCACTTTCCCACTGGTTCACAAGGTGAAGTAGGC
TCACCACAAGACATTCCAGCACTAGATGGAGCAGCTGACAGTCCTGGAGGAGAAAAGCGGATGTGG
CGGCGATTGCCTGATTCCGCAGTCGAGGCATTACGGTGTATTCTTGATGTTAATAGGCTGCCCTCA

FIGURA 19 CONT.

AAATCTCGGAAGGAAGAATTAGCAATTAAGCTTGGGTTGAGCTTTTCACAGGTTCATGGATGGTTT
AAAAATCAGCGGCATCAAGCTCTGAGGAAAGGGTTGGCATGTCCTAAGCGGATACATCCAGTAGCA
TCAAGCTCAAGAGATTTCAATACTGTCGAATCTCCATCGTCTGTGGCGACTTTAACTGAAAGCACA
TTACCGCAAAAAGAGAATGCTCCAAATTCTATTTTAACTGAGAAGTTGGACGAAGTTCAGGTGCGA
CTCCTCGAATTGAAGCAAGCGTTGGACGATTTTACCAAGAATGCAGCAGTGGTTGCGGATGGCGGA
AGTATTGAATTGGATAACAACGTGGAAAACGGAGGAATTGGACTTCCTGAAAATGGAAAACGTGTT
GTCTATGTGCCAGTTGTGGCCGACGTTAGGGAACAATCACCAATGAGTACGCATACTAGCGCATGT
TGA

SEQ ID NO: 241, sequência de ácido nucléico proprietária de Zeama_PHD_HD de Zea mays

MVPTGVVEREDQPEPAAAPPVPVITPKTLRSHSIRTQTTPGDGDAGADHAPGAVAAAAASAGGGGA
ADVMQGNDAGQERRGRRRGRRKKSDVPLDNPTRILKHVKYLLIKMRVHQNLLDAYTGEGWKGQSRE
KIRPEQELQRAKAQILRSKLAIREAIHELDDVGLEGSLDKNAFDSEGRIYHEEIFCAKCKSQEALP
DNDIILCDGACNRGFHQYCLDPPLATEDIPPGDEGWLCPVCDCKMECIEAINSYFGTSFEVENSWE
SFFSNEAGIAAGGGTQEGAGGDWPSSGDEDDEYDPETAEGPSSASESGAQVIGLAVDWPSSDDEDH
DYDPEAAEEPPGASGSGTQTVGLDVDWPSTDDEDDDFDPEGIRKQANSDEESSSDESEEDGEDSGS
VSGFRENEASSGVSDKDEQEPSEKHNLELTAVTNIKKLARKGKRNIISNKGKSEYLESSDSDDSDV
SLFDEFKGGNANLGLRRGDPQAAVSAADEEAVIIAGKRHRKAVDYKRLHDEMYGKAEADNDDDVIS
EDEDWGPERRRRRTIPDDPNSPRSRIRGRKVRTPSGDISKGADADLPSGLPDMPSHFPTGSQGEVG
SPQDIPALDGAADSPGGEKRMWRRLPDSAVEALRCILDVNRLPSKSRKEELAIKLGLSFSQVHGWF
KNQRHQALRKGLACPKRIHPVASSSRDFNTVESPSSVATLTESTLPQKENAPNSILTEKLDEVQVR
LLELKQALDDFTKNAAVVADGGSIELDNNVENGGIGLPENGKRVVYVPVVADVREQSPMSTHTSAC

SEQ ID NO: 242, sequência de ácido nucléico proprietária de Zeama_PHD_HD de Zea mays

ATGCATTCTTCGGAAAATAAACTGGGTTGGAATGAGACCTCTATGGGGTTGGTTCCTGTACCAAAA
AGACCAGCTAGACCTGATGCTTCCCACCAATGCAAATCTGATTCCTTTATGAAGCGATCACCAAGG
AAAGTTAGAAATGCTACTCTGGCAAAAAGCATAAAGAGCAAATACCACTATAGTCCCCTCAAACAG
CGGAAGGGTTCAGATTCTGTTCCTGGGAAAATCGTAACAGGACTAACCGCAAGGAAAAAGAAGAAA
AGGAAAATCCAAATTACAGACGAGGCAACTCGTTTGGAACGAAGAGCGAGATATTTTCTAATCAAG
ATAAAACTGGAGCAGAATTTGCTAGATGCTTACTCTGGAGATGGATGGAATGGGCAAAGCCGAGAG
AAAATAAAGCCAGAGAAGGAACTGCAACGTGCCAGGAAACAAATCATAAAATGCAAAATTGCTATA
CGTGATATCATCCGCCAACTTGATTTGTATACTTCTACTGGGAGTGTTGATGACCCACTAATGCCT
ACAGATCAGTCCACCAATCCCGAACATACCATGTGCTCAACATGCAAGTCTCATGAATCATTTCCC
AGCAATAAAATTATCTTCTGCAAAGGGCCCTGCAAAAGGGCATGTCATGAGAAATGTTTGGAACCT
CCCTTAAACAAAAGCGTACTTCCAACAAGTAGCCATGGGTGGCTTTGCAAATTCTGTTTGTGCAAG
GTGAGGATTTTAGAAACTATTAATGCACATCTAGGGACGAGTTTTACAGTGAAGTGCCACTTTGAA
GATATATTCAAGGAAACTACTGAACTAATAGACTCTGAGGATGCACTAGATGAAGATTGGCTTTCT
GAGTATTCAGGTGATGAGGACTATGATCCTGATGAAAATGAGGCCAGTGGCGACTGTATGGACAGC
GGGGAGAAGATTATGTCTGATGATTCCAATGGTTCAGGAAGCCCCCTTTATTCTCCAAATGACGAT
ATTCCTGACTTCATATCAGCAGACTTAAATGTTGTGGAAGGGTTTTGTCATACCAATTTAGATTTA
GGCATTGATGCCGTTGAAGATGATTTTGCACAGATCCTCACCTACCAAAGGCCAAGAAGAGATGTT
GATTATAGAAGGCTTAACGAGGAAATGTTTGGGAAAATAACCGGGAATGAAGAACAGAGCGAGGAC
GAAGATTGGGGCCATGAAAGGAGAAAGAAAAGGACGCATTCAGGTGTTGCTGGGGATAATTCTGTT
GGTTTCTTGAACGTTATATCTGATGAAAAGAGCCAGAAGAAGGGGAGAAAACTTTTCAGGATTCCT
CCTGCAGCTGTTGAGGTACTTCGCAGAGCTTTTGCTGAAAATGAGCTTCCACCCCGGGATGTTAAA
GAAAATCTCTCAAGAGAATTGGGAATTTCTTTTGAAAAGATTGATAAATGGTTCAAAAATACACGG

FIGURA 19 CONT.

TGTGCTGCTCTCAGAGATCGGAAGGCTGAAGGAAACAGTCATAATACAGCTCCCAGCAAAATCTCA
AAAAATAAAGGAAAAGCTGGAATCTCAGGCAAGGCTGGAAGGAATGGTCATGTTACTGGTCCCTGC
AATAATTTGAGAATAAATGAAGAAAAAACAGGTATATCGGGGAAGTTTGATTCAGGAGACAACTCT
TGTTTGGTTCCCTTCTCTGAAGTCATCAATGTGCCTACGCGATTGCCACACAACTTTGAGATGAGG
AAGATGGAATCAACTAGAAGTCCTGCGAGGCTACACAATAAAGGAGGGTGCCTGTCTGCAACAATC
CAAATTAAGGAGAGTACATTGTTACCCACAGGCAAGCCTTGTTGGCAGTCAGAAATGAGCCATCCA
ACAACTAATGAAGTGAGCACTTCAGCGCAAGCTACTACTTGGATCGACGCGGGGGCGTGTGCCGAG
GAGCAAGAACCTACTCCTTGGGTGGACATCGGGGCCTCAGACTACCAGCCTTTCCTGGATGTGATC
GACGACATGTGCGGACTTGAGTGGAGGCTGCAGAGACTGAAGGAGAACATGCTCTCATCTAGCATA
GACGGCCAAACCGCCGCCAGTGAGAGTGACAAAAGAAACCAGACCGTGGTGCTCGTGCCAACTGCC
GAGCTCAAGGAAAAACTGCCGCATGACGGTTTATTCGGGCATTATTGCCCATAG

SEQ ID NO: 243, sequência de polipeptídeo traduzida proprietária de Zeama_PHD_HD de Zea mays

MHSSENKLGWNETSMGLVPVPKRPARPDASHQCKSDSFMKRSPRKVRNATLAKSIKSKYHYSPLKQ
RKGSDSVPGKIVTGLTARKKKKRKIQITDEATRLERRARYFLIKIKLEQNLLDAYSGDGWNGQSRE
KIKPEKELQRARKQIIKCKIAIRDIIRQLDLYTSTGSVDDPLMPTDQSTNPEHTMCSTCKSHESFP
SNKIIFCKGPCKRACHEKCLEPPLNKSVLPTSSHGWLCKFCLCKVRILETINAHLGTSFTVKCHFE
DIFKETTELIDSEDALDEDWLSEYSGDEDYDPDENEASGDCMDSGEKIMSDDSNGSGSPLYSPNDD
IPDFISADLNVVEGFCHTNLDLGIDAVEDDFAQILTYQRPRRDVDYRRLNEEMFGKITGNEEQSED
EDWGHERRKKRTHSGVAGDNSVGFLNVISDEKSQKKGRKLFRIPPAAVEVLRRAFAENELPPRDVK
ENLSRELGISFEKIDKWFKNTRCAALRDRKAEGNSHNTAPSKISKNKGKAGISGKAGRNGHVTGPC
NNLRINEEKTGISGKFDSGDNSCLVPFSEVINVPTRLPHNFEMRKMESTRSPARLHNKGGCLSATI
QIKESTLLPTGKPCWQSEMSHPTTNEVSTSAQATTWIDAGACAEEQEPTPWVDIGASDYQPFLDVI
DDMCGLEWRLQRLKENMLSSSIDGQTAASESDKRNQTVVLVPTAELKEKLPHDGLFGHYCP

FIGURA 19 CONT.

MAGQPPQGPEDDFLDQFFSLTNSLSAAGRPSGDQPFSLA

---1---

LSLDAASDASGSRGGIGDDAGNAAERDGVQLPGLFPPVF

GGGLQPPHLRPSHPPPQMFHAQQPKQGGPAGGPQPPAPR

PKVRARRGQATDPH**SIAERLRRERIAERMRALQELVPNT**

----2---------- ---------

NKTDRAVMLDEILDYVKFLRLQVKVLSMSRLGGAGAVAQ

-------4----------|---5---

LVSDIPLSVKGEASDGGSKQQIWEKWSTDGTEKQVAKLM

DEDIGAAMQFLQSKALCMMPVSLAMAIYDTQHSQDGQPV

--6-------- ------

KPEPNNTA

FIGURA 20

CLUSTAL 2.0.3 **alinhamento de sequência múltiplo**

```
AT2G24260                       ----------MMNSSLLTPSSSSSSHIQTPS-
AT4G30980                       ----------MNSSSLLTPSSSPSPHLQSP--
AT5G58010                       ----------MENG-----NGEGKGEFINQ--
TA12416                         ----------MQPCSSGGEMQGISSLLSNGSH
TA5414                          ----------MQPCSSGGEMQGISSLLSNGSH
TC205173                        ----------MQPCSSGGEMQGISSLLSNGSH
GSVIVT00017237001               ----------
TC7102                          ----------MQAMN-----------------
TA34455                         ----------MQAMNSLLSQQQQSQMSLQDLQ
TC7103                          ----------MQAMNSQMSLQEENGASSGQ--
TA37666                         ----------MQAMNSFQSTGE-NGASSGE--
TC67603                         ----------MQPCSREMQAMNSLLNPTQQIPLQDLQIN
Os06g08500                      ----------MQPSSRDTV------AGG----
Os02g55250                      ----------MQPNARQMR------GGG----
Os09g25040                      ----------MAAAAAAGQISLDDLR------NGGGVAA
TA2544                          ----------MQPSSKETERMAQSHAALQLEL
DV982110                        ----------
WO051050seqID516                ----------MSLLPPGIGGQGLNSQENDPYP
AT4G02590                       ----------MASNN--PHD-----------
TC10015_part                    ----------
AT1G03040                       ----------MANNNNIPHD-----------
TA13791                         ----------MANNP----------------
TC60118                         ----------MANNP----------------
TC60119                         ----------MANNT----------------
TC61833                         ----------MANNP----------------
TA3392                          ----------MANNP----------------
WO051050seqID1671               ----------MANNP----------------
Pt_TC63334                      ----------MANNP----------------
TA3263                          ----------MANNP----------------
Pt_scaff_II.416                 ----------MANNP----------------
TA2825                          ----------MANNP----------------
TC207545                        ----------MANNP----------------
TC229602                        ----------
TA46156                         ----------MASNP----------------
```

FIGURA 21

```
AK247217        ----------MANNP----------
TA30646         ----------MANNP----------
TC8633          ----------MANNP----------
TC15501         ----------MANNP----------
TA1140          ----------MANNN----------
TA3490_part     ----------MSTDP----------
CK293938        ----------MATDP----------
TC12771         ----------MATNQ----------
TC172581        ----------MATNQ----------
TA28621         ----------MAANQ----------
TA21665         ----------MAANQ----------
TA46194         ----------MADNP----------
DY268946        ----------MATNP----------
CO123623        ----------
Pt_scaff_70.65  ----------MAGNP----------
Pt_scaff_XIII.403 ----------MAGNP----------
TA1616          ----------MAGNP----------
TA55042         ----------MAGNS----------
WO051050seqID77 ----------MAGSN----------
TC3698          ----------MASNN--NP------
Os03g58330      ----------MAGQQ----------
TA139285        ----------MAGQP----------
TabHLH11        ----------MAGQ-----------
TC104646        ----------MAGQP----------
TA126400        ----------MAGQP----------
Os07g08440      ----------MAGQQ----------
Os02g35660      ----------MGGFAYPFTPSPAWSRDAVFAGSPWAAGG
DT543504        ----------MIS---
Pt_scaff28.86   ----------MISSNQ
GSVIVT00016367001 MDEYLDHLFSSTSWSDVNVKEGSSWICGEPSQTNGMLSDSIGIYEGDKKNSPVGMTNSNL
TC19278         ----------
TA14134         ----------
TC68930         ----------
TC140470        -MNYSDGSFFPSWPGNSASENYSFVDGSVESYTEEGSMPTTGYFR-ARSNQNLTFDEHEQ
TC253044        ----------
Os06g09370      -MDYSAGSYM--WPGNSGSENYNFVDGSSESYAEEGSLPPSGYFMGAGSDRSLKITENER
```

FIGURA 21

```
AT2G24260          ----------TTFDHEDFLDQIFSSAPWP----------
AT4G30980          ----------ATFDHDDFLHHIFSSTPWPS----------
AT5G58010          ----------NNDFFLDSMSMLSSLP----------
TA12416            ----------EQQSQIHQSATFDPT-SQDDFLEQMLSSLPSCS--WTDLKSPWGVVDL
TA5414             ----------EQQSQIHQSATFDPT-SQDDFLEQMLSSLPSCS--WTDLKSPWGVVDL
TC205173           ----------
GSVIVT00017237001  ----------MLSTLPSWSDLPANPKSPWELNAS
TC7102             ----------QYDPTSSHDDFLDQILSSVPSSSLCWPDLSKSW-----
TA34455            NGGNGGSTGGVGGLGQHNMGHFHFDPTSSHDDFLEQILSSVPSSSP-WPDLSKSW-----
TC7103             ----------NMSHSHFDPTSSHDDFLQQILSSVPSSSP-WPDLSSGG-----
TA37666            ----------HISHSHFDPSSSHDDFLQQILSSVPSSSP-WPEISGDG-----
TC67603            -GNRHHHQQIHVPSSSQFQYPTPTSTTHHDDSFLEQILSSTTSFP--WSDETGPP-----
Os06g08500         ----------GGEG-TQDDFFDQMLSTLP-SAWADLGGG--GGGAAG
Os02g55250         ----------GGGGGGQDDFFDQMLSTLP-AVWSELGSGKPAWDLTA
Os09g25040         ----------NAGGGGGVHDDFLDQMLSSLPPSAWPDLAAGKAAEDDAE
TA2544             ----------QNGGGGAQNDDFFEQTMSGFSTAAWAP-ELGTPRSLFGA
DV982110           ----------
WO051050seqID516   ----------YDDSQFLANRLRQHQLSGNSSMNQTPNQASQGINETNT
AT4G02590          ----------NLSDQTPSDDFFEQILGLPNFSASSAAGLSGVDGGLGG
TC10015_part       ----------HASETPSDDFFEQILGLPNFSASSS------DGGLGG
AT1G03040          ----------SISDPSPTDDFFEQILGLSNFSGSSSGSG----LSGIGG
TA13791            ----------NES-PADDFLEQILGLSHFAPTET-GLAGPDGRLSG
TC60118            ----------NES-PADDFLEQILGLSHFAPTET-GLAGPDGRLSG
TC60119            ----------NEAPAADDFLEQILGLPNFAPSET-GLAGSDAGLAA
TC61833            ----------NEA-SADDFLEQILGFPNFAPSET-GLAGPDGGLTG
TA3392             ----------NEASSTDDFLEQILGIPNFGSAES-GLAASDGGLA-
WO051050seqID1671  ----------SDGASDEFLEHLLGLPNFAAAAEAGLAPGDGTGLS
Pt_TC63334         ----------TEPPTDDFLQEILGMPNFASAEAG-LVGADAGLAG
TA3263             ----------TEPPADDFLQEILGMPNFASAEAG-LVGADAGLAG
Pt_scaff_II.416    ----------TEPPTDDFLQEILGMPNFASAEAG-LVGADAGLAG
TA2825             ----------TEPPADDFLQEILGLPNFASADAAGLVGADGALA-
TC207545           ----------
TC229602           ----------
TA46156            ----------SEAPADDFLEQILGIPTYPAADP----------NLAA
AK247217           ----------SEGPSDDFFDQILGFPAYNGAEP-NLAGNDAGAIP
```

FIGURA 21 CONT.

```
TA30646          ----------SEGPSDDFFDQIMGFPAYNGAET-NLAGNDAGAIP
TC8633           ----------SEGP-DDFFDQILGFPAYNGAET-NLAGSDAGAIP
TC15501          ----------NPTEGLTDDFLDQILSYG----AHEGNLAGNDVNLAG
TA1140           ----------PEGY-GDDFLEQILAIPSYN----------
TA3490_part      ----------PEGY-GDDFLEQILAIPSYN----------
CK293938         ----------PEGY-ADDFLEQILAIPSY----------
TC12771          ----------PGGY-ADDFLEQILAIPSY----------
TC172581         ----------PEGY-ADDFLEQILAIPPY----------
TA28621          ----------PEAY-ADDFLEQILAIPSY----------
TA21665          ----------PEVYAADDFLEQILAIPSY----------
TA46194          ----------PEGY-ADDFLEQILAIPSYP----------
DY268946         ----------MFLQLGSGGPSPGAG----------
CO123623         ----------DSG----------
Pt_scaff_70.65   ----------PPEGLGDDFFEQILAGQPPGYGG----------EA
Pt_scaff_XIII.403 ----------PPEGLGDDFLEQILAAQPPGYGGG----------GEV
TA1616           ----------PPEGLGDDFFEQILAVQP-GYGGGDG----GGGGDV
TA55042          ----------GSEGLGDDFFEQILAVPEAGTVG----------
WO051050seqID77  ----------PQEGLGDEFFEQILAVPPGAYSGGS----------TPM
TC3698           ----------QPPSDGLNDDFFDQIFSMPSAFAASS----------
Os03g58330       ----------PQQQGPPEDDFFDQFFSLTS--SF----------
TA139285         ----------PPQG--PEDDFFDQFFSMTAGGSY----------
TabHLH11         ----------PPQG--PEDDFLDQFFSLTN----------
TC104646         ----------PPPGG-SEDDFLEHFFAFPSAASA----------
TA126400         ----------PPSGA-SEDDFLEHFFAFPSAASAG----------
Os07g08440       ----------PPTGG-AEDDFFDHFFSIPSAAAA----------
Os02g35660       ----------VSSLADALVSYGAVDDEEAAFLGKTAASSPSTARLHEQQQLLLE
DT543504         ----------
Pt_scaff28.86    SVESLATQDTSSVVLGSESDYAVDKVLISEQARLQNDCQNCNGNPSPDGMARGNLKFGNT
GSVIVT00016367001 MIEGLVTQDTSSIVLGGESDHGLGKGLLLEEAPRQQDLQNTEGNSSLNGAVNGSSEVGYV
TC19278          ----------
TA14134          ----------
TC68930          ----------
TC140470         NPAMLANGCLPYNTQTDLLSGEILSDDKPSNSLVELSQLQNNVCLQNNLIPPGTLQCNST
TC253044         ----------
Os06g09370       NPTMLANGCLPYNTQAHPLSGQILPKGELPNNLLDLQQLQNSSNLRSNSIPPGVLQCNST
```

FIGURA 21 CONT.

```
AT2G24260          ----------------------------------SVVDDAHPLPSD-----GFHG
AT4G30980          ----------------------------------SVLDDTPPPTSDCAPVTGFHH
AT5G58010          ----------------------------------PCWDPSLPPPPP---------
TA12416            NPNNNNNNINNKQPRDLSDETAPSTTQENNVPAGFQFDESMILASKMRQHQISGNTGS-G
TA5414             NPNNNNNNINNKQPRDLSDETAPSTTQENNVPAGFQFDESMILASKMRQHQISGNTGS-G
TC205173           ----------------------------------------------------------
GSVIVT00017237001  NP----ISMPSNKSRDLSDDTTPSNPDN----VQFAFDESAMLASKLRQHQIS------G
TC7102             DPHHHHLSPPLPPNPSSGDDHQLPPSNPLQ-HFQYDDQSSSFLAAKLRQHQITGGGGGGG
TA34455            DPHHHLSSP--PHNPSSGEDQ--PPSNPLHSQFHYDDQASSLLASKLRQHQITGGG----
TC7103             DGH-------------------------FDD-QSIFVASKLRQNQITGGGG---
TA37666            HPYN--------------------------FDDHQSTLLASKLRQHQINGGSS---
TC67603            NPEDNNNAG----------------------FNYDEMVLASKLRQHQINGGAAG-N
Os06g08500         KSPWEVDPAA---------------------AAAASQVFDESALLASRLRHHQIGGAGGGGG
Os02g55250         GAVGGGGGAS----------------------DDHSAAAFDDSALLASRLRQHQID----GGG
Os09g25040         G-MHHHHHQQ----------------------QQQFGGPYDESAMLASRLRQHQISGGGGGGG
TA2544             RSSEESPDDE----------------------GLNYAPPYGGSSLLASRHRMHLGT----SPA
DV982110           ----------------------------------------------------------
WO051050seqID516   SPGRSMVLQLS----------------------TGSASASQLLASMGNSPRGTAVMTQSPSTG
AT4G02590          G-------------------------------APPMMLQLGSGEEG-------
TC10015_part       GG-------------------------------APPMMLQLGSGEEG-------
AT1G03040          VG---------------------------PPPMMLQLGSGNEGN-------
TA13791            NATT-------------------------AGAPMLLQLSSGGGT-------
TC60118            NATT-------------------------AGAPMLLQLSSGGGT-------
TC60119            TAAG-------------------------AGAPMFLQLSSGDGA-------
TC61833            TTAG-------------------------AGTPMFLQLSSGDGA-------
TA3392             ----------------------------AGSPMMLQLSSGDGS-------
WO051050seqID1671  VT---------------------------AATPMMLQLGSGDGSGGHGHGH
Pt_TC63334         AAA--------------------------AQASMMLQLSSGDGS-------
TA3263             AAA--------------------------AQASMMLQLSSGDGS-------
Pt_scaff_II.416    TAS--------------------------VQAPMMLQLSSGDGS-------
TA2825             ----------------------------TPMMLQLSSGDGS-------
TC207545           ----------------------------MMLQLNSGDAA-------
TC229602           ----------------------------------------------------------
TA46156            NDVN-------------------------LAAPMVLQLGSGEGS-------
AK247217           ----------------------------PAMMLQLNSGDGS-------
```

FIGURA 21 CONT.

```
TA30646          ----------PAMMLQLNSGDGS-------
TC8633           ----------PAMLLQLNSGDAS-------
TC15501          TTT-------PGGSMALQLSSSGGD-------
TA1140           --IT------GCLPVNTADSTGSET-------
TA3490_part      --IA------GCLPGNTTDANASET-------
CK293938         ----------AGLPPVADVGTSSET-------
TC12771          ----------AGLPPVADVGTSSET-------
TC172581         ----------SGLP-VADVGTPSET-------
TA28621          ----------SGLP-VADVGTPSET-------
TA21665          ----------ASLP-VTDLTAGASS-------
TA46194          ----------GHDPNMVGTGSTMVL-------
DY268946         ----------GGATSAADIRSLA
CO123623         ----------GGG-----FR---
Pt_scaff_70.65   VGST------SLPMMGLQLGSSAANVGLMRSS
Pt_scaff_XIII.403 MGST------SLPMMGLQLGSCAGNVGLLRSN
TA1616           VGST------MPMMGLQLGSASSGGGGLRT-
TA55042          ----------MLQLGSTTGAFRGAS--
WO051050seqID77  VLQL------GSGRGGVEDGGGRGGGGGGLRG-
TC3698           ----------ASTDATEALNYADSS-------
Os03g58330       ----------PGAAPGG-RAAGDQ--------
TA139285         ----------PGATAGGGRAPGDQ--------
TabHLH11         ----------SLSAAG-RPSGDQ--------
TC104646         ----------AAGGHAGAGAGGDH-------
TA126400         ----------AAGGHAGAGAVGGDH-------
Os07g08440       ----------GAGGVGGFGSGDHH--------
Os02g35660       AELLRHGDGLG-------FAAMDDDGGAAMLGALEPCAMPLTDSGGPPVICSSS
DT543504         ----------NGGESSEFRRSFT----
Pt_scaff28.86    GLQCNGILPTLSSLNY--PNQLPIVGDLT-SYLSFSEASNAGCNGREQSEYLRSLK----
GSVIVT00016367001 GLQLNTPTSTPCSLDLGSPKQLSLIGGMSRSSRPFTELDHVGCNGNEPSDFQRSVG----
TC19278          ----------MDDIFDQLLSSSWE----
TA14134          ----------MLARENRSGDTEHD----
TC68930          ----------MLLCSTIQVIQ-------
TC140470         PGTFDLQLDTPGLLELPHALSSSIESNGSEVSAFLADVQAVSSASTLCSTFQNVPSYMEP
TC253044         ----------MLLCSTIQVIQ-------
Os06g09370       SGTFDAKLDTPGLAELPHALSSSIDSNGSDISAFLADVHAVSSAPTLCSAFQNVSSFMEP
```

FIGURA 21 CONT.

```
AT2G24260          HDVDSRNQPIMMMPLNDG----------
AT4G30980          HDADSRNQ-ITMIPLSHN----------
AT5G58010          ----------
TA12416            ----------
NNNSAAAKFMMQQQQQQ----------IMMAAARGGIG----------LPLSLGNG-----
NNNSAAAKFMMQQQQQQ----------IMMAAARGGIG----------LPLSLGNG-----
TC205173           ----------MLQHQLLR----------DSGLLN----------MPLSLPG------
GSVIVT00017237001  NSSAAKSALMLQQQLLLSRGVAMGRSPSNGSGAGESGLLQ----------LPLSLSNGDSCLV
TC7102             DAVAAAKALLLQQQLLLS----------RTLAGNGLR----------SPTGASGDNGFL-
TA34455            -AAAAAKALMLQQQLLLS----------RTLAGNGLR----------SPNGASGDNGLLS
TC7103             -GGAAAKALMLQQQLLLS----------RGLLAG----------
TA37666            -AAAAAKALMLQQQLLLS----------RGIAGN----------
TC67603            PFAAMKMMLMMQQQQQQQ----------QMMLAGRPT----------VASAAAGG-----
Os06g08500         EKP----------VMLQLSELHR----------QAG---GGEEDGSGA----------FSPLPLFTDRTNVP
Os02g55250         DKP----------IMLQLSDLHR----------HHGLAAGDDSGGAAG----------FLPLSLFADRS---
Os09g25040         GGAAAVKQMVLQQLADLRQGHHMM-LQGLGGRSPAGGGGGGDGGLLLPLTLGSGGS---
TA2544             DSS----------MFLQLGGQILLHP---TAG-ASGGESGGFLR----------LPLSLGSGGSGVS
DV982110           ----------MLPLPLSLG----------QAGKSGDMN----------
WO0510505eqID516   GSCNGSDGGMLPLPLSLG----------QAGKSGDIN----------
AT4G02590          ------SHMGGLGGSGP----------TGFHNQ----------MFPLGLSLDQGKG-
TC10015_part       ------GHMGGLGGG----------SGFHNQ----------MFPLGLSLEQGKG-
AT1G03040          -----HNHMGAIGGGGP----------VGFHNQ----------MFPLGLSLDQGKG-
TA13791            ------GHIGAIGGGGG----------GAFHGQ----------VFPLGLSLEQGK--
TC60118            ------GHIGAIGGGGG----------GAFHGQ----------VFPLGLSLEQGK--
TC60119            ------AHIGGIGGGGG----------GAFHGQ----------VFPLGLSLEQGK--
TC61833            ------SHLGGIGGGGG----------GAFPGQ----------VFPLGLSLDQGKS-
TA3392             ------SHISALGGGVS----------SGYHGQ----------VFPLGLSLEQGK--
WO0510505eqID1671  GHGHGQGHLSALGGGGAGAGGGGG-----GGGGGYHGA----------VFPLGLSLEQGK--
Pt_TC63334         ------GHISDLGG-APGGGS----------AGFHG----------FPLGLSLEQGK--
TA3263             ------GHISDLGG-APGGGS----------AGFHG----------FPLGLSLEQGK--
Pt_scaff_II.416    ------GHISALGG-APGGGG----------AGFHG----------FPLGLSLEQGK--
TA2825             ------NHITALGGGGGGGGG----------AGFHG----------FPLGLSLEQGK--
TC207545           ------THLA----------SFHAP----------PYQLGLSLDQGE--
TC229602           ----------
TA46156            ------GHIAG----------GYQGT----------MFPLGLRLEQGK--
AK247217           ------SQFTGVGLGVGLG----------GGG-FHGHGG----GGSFPLGLSLEQGK--
```

FIGURA 21 CONT.

```
TA30646          -----GQFTGVGLGVGLG----------GGG-FHGHGG----GASFPLGLSLEQGK--
TC8633           -----GQFSGMGLGVGLG----------GGGGFHHPSQ----SGSFPLGLSLEGGK--
TC15501          -------ITGSGGYHHHHQHQHH-----HQLSGGGGG----------FPLGLSLEQSGK-
TA1140           ------VSVHHQQQ---------------QPQQQLQ----PQKFPLGLSLDNGRET
TA3490_part      -----VSVHRQQQ---------------QQQ----------PVFPLGLSLDNGRET
CK293938         ------TSFTSAS---------------AAGLQ----QPLFPLGLSLENGRDD
TC12771          ------TSFTSSSVA-------------AAGLQ----QPLFPLGLSLDNGRDD
TC172581         ------TSFTSASAVSHLN---------SAAAAGLQ----QPLFPLGLSLDNGRDD
TA28621          -----TSFTSASAVSHLN----------SAAAAGLQ----QPLFPLGLSLDNGRDD
TA21665          ------ENSTSGVS--------------QLQQ----QPLFPLGLSLDNGFAD
TA46194          ------QLSSGDGSG-------------HVAGGGFQ----GPVFPLGLSLESGIPA
DY268946         MG--ISMGMMPLGLNLEHS----------FLRQHE-----DHNSNHNNNNNTNA-
CO123623         -----GIGMMPLGLNLEHG----------FLR-HE-----DGVVVDNNNNNNASC-
Pt_scaff_70.65   AN--NNMGMMPLGLNLEHHG----------FLRQQQ-----DDGSSSLDTNNSNN-
Pt_scaff_XIII.403 AN--NNMGMMPLGLNLEHHG----------FLRQQQ-----DDSGSSLDTNNSNN-
TA1616           ---NNMGMMPLGLNLE-------------FLRQQ------EDGSGALDNNNHTN-
TA55042          -------GLMPLGLNLEQAA---------FLRHQVN----VDDDVVHVNVDDATIH
WO051050seqID77  -----MGVMMPLGLNLEQGG---------LFRHED-----VENSTSASSTTSAIN-
TC3698           -------TTGVTP----------------------------MFSLGLSLDQQPK-
Os03g58330       ------------------------------------------PFSLALSLDAAAAA
TA139285         ------------------------------------------PFSLALSLD-AAAA
TabHLH11         ------------------------------------------PFSLALSLDAASDA
TC104646         ------------------------------------------PFPLALSLDAAAE-
TA126400         ------------------------------------------PFPLALSLDAAAE-
Os07g08440       ------------------------------------------PFPLALSLDAEGAG
Os02g35660       SNDSSGSEHSAAMPAGGGFLVGEQQQHVPPAAYAAGGVLPSMAAGEETPQSFGFGSLFNG
DT543504         ----------------------GLETLSPIPQLWHLQPYDSVSSLPTLVGQTIVD
Pt_scaff28.86    ----------------------NLQNLSSIPQLWPSQSYEGVSSLPPLMGQDRIE
GSVIVT00016367001 ----------------------NFQALPPIPQLWSQPSYGGGSSLSPVMGEYKMQ
TC19278          ------------------------------------------
TA14134          ----------------------DINGAD------------RSSLDSSAG-----
TC68930          ---------------------DLQGTL--------------LTSYTGPQG-----
TC140470         VSLEAFSFQGMQNAAMFNNASHSNGNLSVFDEATMASLHDSKEFLNG-SISSLGTGQQSQ
TC253044         ---------------------MGNLSVFDEATMASLHDSKEFLSG-SISSFGTAEQSQ
Os06g09370       VNLDAFGFQGAQNVAMLNKTSLPNGNPSLFDNAAIASLHDSKEFLNGGSIPSFGTVLQAL
```

FIGURA 21 CONT.

```
AT2G24260          ----------SSVHALYNGFSVAGSLPN----------FQIP
AT4G30980          ----------HPNDALFNGFSTG-SLP----------FHLP
AT5G58010          ----------PPQSLFHALAVDAPFPDQ----------FHHP
TA12416      ----------GDNDIVD-VSSFKS-QQGGDGSVQALYNGFT-GSLHG----STQPQHFHHL
TA5414       ----------GDNDIVD-ASSFKS-QQGGDGSVQALYNGFT-GSLHG----STQPQHFHHL
TC205173     ----------NDVVVDASTFESPNPSGKASVQTLYNGFA-GSLHGVGQSSNQTQHFQHP
GSV1VT00017237001  DR----------SQNDVVDGSSSFKSPNQGGDGSVQALYNGFA-GALHGSGQASNQAQNFHHP
TC7102       ----------NMLGNGDQNDSVGNPANDSSVQALFNGFT-GSLGQ---NSSQPQHFLHP
TA34455      -L----------PLNLSNNGDQNDGVANPANDNSVQALFNGFT-GSLGQ---TSNQPQHFHHP
TC7103       ----------NGDQNDDVGNPENEISVQALYNGFA-GSLGQ---TSNQPQHFHHA
TA37666      ----------GGDQNDDGLNPGNDISVQALYNGFA-GSLGQ---TSNQSQHFHHS
TC67603      ----------GITTNHQGGEGSMQALYNVFGDGSLHG----TMQAKSFVGA
Os06g08500   P----------REEMEGGFKSP--NAAAGGEHALFNGFGVHGGSG-----GAGQ--PPF
Os02g55250   ----------QDDIDAAFKSP--NGARG-DHALYNGFGAAGMHG-----AAAMQPPPF
Os09g25040   ----------GGDVQALLKAAAANSAGGDAGGVYGGGFAGSLH-----QQQQHFQFH
TA2544       TS----------AIFGDRSRGEIDPP-FESSNPTEAEVLYNDGFTGSLP----RAVQASLHQ
DV982110     ----------EPGSREEIEASFKSVNNARDSNLGGLFQPFAVSPRGVR----PTGQNFHAQ
WO051050seqID516   ----------ESRSRDEIEASFKSANHARDSNLGGLFPPFAASPRGVR----PTGQNFHTQ
AT4G02590    ----------PGFLRPEGG-HGSGKRFSDDVVDNRCSSMKP----------
TC10015_part ----------QGFLRPEGGSLGTGKRFSDDV--------MKP----------
AT1G03040    ----------HGFLKPDET----GKRFQDDVLDNRCSSMKP----------
TA13791      ----------GGFLKPEEASG-SSKRFRNEVVDGRAFSVKN----------
TC60118      ----------GGFLKPEEASG-SSKRFRNEVVDGRAFSVKN----------
TC60119      ----------GGFLKPEEASG-SGKRFRNGVVDDRASSVKN----------
TC61833      ----------GGFLKPEEGSGGSSRRFRDEVVDGRASSVKN----------
TA3392       ----------GGFLKPEEASG-SGKRFP------EEHAIKN----------
WO051050seqID1671  ----------GGFLKPEEASG-SGKRYPEEVVDGRASTVKNLHSPLPNFPDV
Pt_TC63334   ----------GGFLKPEEASG-SGKRFRDEIVDGRA---KN----------
TA3263       ----------GGFLKPEEASG-SGKRFRDEIVDGRA---KN----------
Pt_scaff_II.416    ----------GGFLKPEEASG-SGNRFRDDIVDGRV---RN----------
TA2825       ----------GGFLKPEEASG-SGKRFRDDVVDGRANTVKN----------
TC207545     ----------GPFLTPEDASG-SGKRFRDDAVDTRP---NN----------
TC229602     ----------MKPDEASA-SGKRFRDDVVDNRAKH----------
TA46156      ----------SSFLKPEDASG-SGKRFREEVIDGRASTVKN----------
AK247217     ----------GGFLKMDDV-SAPGRRFRDDVVDSRASSSVKP----------
```

FIGURA 21 CONT.

```
TA30646                 ----------GGFLKMDDV-SAPGRRFRDDVVDSRASSSVKP----------
TC8633                  ----------SGFMKMDDVPAVPGRRFRDDVVDSRASSSVKP----------
TC15501                 ----------GGFFKPDEASG-SGKRFRDDFVDLKAVSSTNNDNNDRGDSVH
TA1140                  IGGA-FAGQ-----QQQLQQQQQRERGGSMNMSGLFPQFENLQSHSLLHTVP---
TA3490_part             IG-A-FAG----------QQQQRERGGSMNMTGLFPSFENLQSHSLLHSVP---
CK293938                VS---DAG----------AYAVKHEREG-INIGNLYAGLEHLQSHAVRHAVPPVHHI---
TC12771                 VS---DAG----------AYAVKHEREG-INIGNLYAGLEHLQSHAVRHAVPPVHHI---
TC172581                VG---DAG----------PYAVKHERDG-MNIGNLYAGLEHLQSHAVRHSVPSVHHV---
TA28621                 VS---EAG----------AYAVKHERDG-MNIGNLYAGLEHLQSHAVRHSVPSIHHV---
TA21665                 AN---NTG----------GFQVKTEREA-MNMGNLYPGLEHLQSHAVCLSVPQVHQV---
TA46194                 TQVAPGSGERFRDDVDARGSSGRSERES-VHLDGLFPAYGHVQSLSVRPAVPQVH-----
DY268946                --------SSSSPNNSSATSGINERDSVHMPSLFPAFGQLQ--SQSAR--PPLPRPPQ---
CO123623                --------SAAS-----AVSGISERDSMHMVSLFPPFGQMQ--TQQIRASPPPPQPPPQ--
Pt_scaff_70.65          --------INNASPFSITSAGFLGRDSVHMTSLFPTFGQLQ--IHSSIRSAPPPLGPPQ--
Pt_scaff_XIII.403       --------INNTPP-SIKSARILGRDSVHMTSLFPTFGHLQ--IHSSIRPTPPPPGPPQ--
TA1616                  --------NNNNNVSSSTTSVINDRDSVHMASLFPTFGQLQ--TQT-LR--PPP--PPH--
TA55042                 QHHLTLHNNNNSSSPSSTAPITDRDSMHMRGLFSAFGQLHTPIRPTLPLPPPPRQPQLHL
WO051050seqID77         ----------MEREAGVHHGNNNMTS-CLFPAFGQFQ--THQSVQPPPPHPPPPQ--
TC3698                  ----------SDSSQSEREHPVQFACLFPQYGHNNQVHQTRPNLSPP-----
Os03g58330              EASGSGKRLGVGD--DAEGGGSKADRETVQLTGLFPPVFGGGG-VQPPNLRPT-------
TA139285                EASGSGK----------HADGG--KADREAIQLPGLFPPAFGGG--VQPPHLRAT-------
TabHLH11                SGSRGGIG----------DDAGN--AAERDGVQLPGLFPPVFGGG--LQPPHLRPSH------
TC104646                ----------PKPDRDPVQLAGLFPPVFAGAGGVQQPHLRGP-------
TA126400                ----------AKPDRDPVQLAGLFPPVFAGAGGVHQPHLRGP-------
Os07g08440              AARRLLDG----------GHDGGRTDRDPVQLAGLFAPVFGAAAGVQPPHLRAP-------
Os02g35660              ----DLLQEATVSKYHHHQQQQQLGVVPSSQPHHLNDDIDFNTGKLMSFASGQQHVTPSI
DT543504                DE----GNIN----------RFIEIDEILQPENLSASINSKGQQDMQNSFCSSFPADHP---
Pt_scaff28.86           GSGLRGGNLDDDMHIMGKGYMGMDEILRLDKLSASPTTEGKEDLQSCPFSSGIAEPN---
GSVIVT00016367001       GFGLQGEYVDNEMDIMRNRYVG-DEILQLDNISSAIPIKGKEEQHTHPFSPSAVGPH---
TC19278                 ----------
TA14134                 ----------GPTNCLPGDSMGTLQASSRVSTMSVTPSSHMSPNLE---
TC68930                 ----------GQT-VLPR-TPGAQSHQQNLSTQTIQSGEGTSLQQY---
TC140470                LAGGGLKAEQQ-----EQNTMCNIPLPSFVSASQMAVSEAQGALIPSKTTSITHNNK---
TC253044                LAGSGLKAEQQ-----EQNAMCNIPLP-FASGSQMAVSEAQGAMIPSKISSTIHNNK---
Os06g09370              GAG-GLKAAQQ-----EQN-IRNIPLPTFTSGSHLAVTDAQGPPLPSKIPPLIHDHN---
```

FIGURA 21 CONT.

```
AT2G24260          QG----------SGGGLMNQQGQTQTQTQPQASASTATGGTV---------
AT4G30980          QG----------SGG---------QTQTQSQATASATTGGAT---------
AT5G58010          QE----------SGGPTMG----SQEGLQPQGTVSTTS-------------
TA12416            QGGS--------MPGQTFGAP--GPVMNQTQAQASGSTGGGGGGG-------
TA5414             QGGS--------MPGQTFGAP--GPVMNQTQAQASGSTGGGG----------
TC205173           QGSSN-------PMQGQNFGAVPAGGGSATNQAPASGAAAGGA----------
GSVIVT00017237001  QGGS--------MQAQNYG-----------APAT-------------------
TC7102             QGGR--------MQAQSFGAPATAPAMNQTPAASG-SAGGG---T--------
TA34455            QGGS--------MQSQSFG----APAMNQTPAASG-SAGGGGG-S--------
TC7103             QGGS--------MQAQSFGAPASS-TMNQTPAASGGSAGGG------------
TA37666            ----------QAQSFGAPAASLTMNQTPAASG-SAGGA---------------
TC67603            NSATE-------MTQSQGSG-----PTAGEAVTAPE-----------------
Os06g08500         GQ----------------------------------------------------
Os02g55250         GQGG--------SMPAQSFGGGAAASGGGGGGSASAAAAAGASS-G-------
Os09g25040         PQTAP-------TIPTQSFGGGGGGGGGGTASGGGAAQPQAGAA-G-------
TA2544             QLLR--------QPQNYGAATLTG------QALAITATASGT-A---------
DV982110           PGQ---------VPLQGYGGMPQPQHQNPPPTGVG------------------
WO051050seqID516   SGQ---------VPMQVYGGMPQPQHQNPSPTGVG------------------
AT4G02590          ----------------------VFHGQPMQQ--PPPSAPH-------------
TC10015_part       ----------------------VFHGQPMQQQ-PAPAAPH-------------
AT1G03040          ----------------------IFHGQPMSQ--PAPPMPH-------------
TA13791            ----------------------VFHGQPVPA--TVAAGPH-------------
TC60118            ----------------------VFHGQPVPA--TVAAGPH-------------
TC60119            ----------------------VFHGQPMQA--TVSAAPH-------------
TC61833            ----------------------VFHGSPMPA--TVAPSPH-------------
TA3392             ----------------------VFHGQPLPS--PVPAAPH-------------
WO051050seqID1671  SQMASFRN----------------RDVFHGQPVPN--PVPAAPH----------
Pt_TC63334         ----------------------VFHGQPMPT--TVAIAPH-------------
TA3263             ----------------------VFHGQPMPT--TVAIAPH-------------
Pt_scaff_II.416    ----------------------VFHGQPMPT--TVTAATH-------------
TA2825             ----------------------VFHGQPMPT--TMAAAPH-------------
TC207545           ----------------------VFDGQPMPT--TVPTAPH-------------
TC229602           ----------------------VFHGQPMPT--TMPAAPH-------------
TA46156            ----------------------AFHGQPMQA--TVAAAPH-------------
AK247217           ----------------------GFHGQPMPS------MPH-------------
```

FIGURA 21 CONT.

```
TA30646          ----------------------------------GFHGQPMPS------MPH---------
TC8633           ----------------------------------GFHGQPMPS------MPH---------
TC15501          LNNLFPAFGHVQ--QQQQAHSVRPPSHQNQINQPFLGHPTPG--AVAVVPH---------
TA1140           ----------------------------------QGFQGQPTSS--TAVTVPH---------
TA3490_part      ----------------------------------QAFQGQSTTS--TAVTVPH---------
CK293938         ----------------------------------QSFQGQPTTS--TTMTVPH---------
TC12771          ----------------------------------QPFQGQPTTS--TTVTVPH---------
TC172581         ----------------------------------QPFQGPPTTS--TTVTVPH---------
TA28621          ----------------------------------QPFQGPPTTS--TTVTVPH---------
TA21665          ----------------------------------QPFQGHPTSS--AIVTIPH---------
TA46194          ----------------------------------QAFHGQPTPV--PITAAPH---------
DY268946         ----------------------------------VQQFQSQP------AAAPQ---------
CO123623         ----------------------------------LHQPFHSQPTSGPV--AAAPH---------
Pt_scaff_70.65   ----------------------------------IHQ-FNSQPTSGAV--SAVPQ---------
Pt_scaff_XIII.403----------------------------------IHQ-INSHPNPGAV--SAVPQ---------
TA1616           ----------------------------------LHQPFHGQPTPGVV--SAASQ---------
TA55042          H---------------------------------HHNQFQGQAAAAPASMAAMPQ---------
WO051050seqID77  ----------------------------------LHPAFLNQPAVG------SPN---------
TC3698           ----------------------------------QVLHGQAMQS--SMAATPQ---------
Os03g58330       ----------------------------------PPTQVFHPQQSKQGGAAVGP----------
TA139285         ----------------------------------PPTQVFHAQQPKQGGAAVGP----------
TabHLH11         ----------------------------------PPPQMFHAQQPKQGGPAGGP----------
TC104646         ----------------------------------PPPQMFQAQ-PKPGEGGMAP----------
TA126400         ----------------------------------PPPQMFQAQ-PKPGEGGMAP----------
Os07g08440       ----------------------------------PPPQVFHAQ-PKPGEGAMAAPQP--------
Os02g35660       DSLQIDQKEFSSGLHHLNLSSLISGPLASFNATQSHRQPAEACGGKNGGAAPFVNLSEVL
DT543504         ----------------------------------ITKTMIGLPSLVQGASPNLN----------
Pt_scaff28.86    ----------------------------------VNMSMNQLSSMPQTTSAAPV----------
GSVIVT00016367001----------------------------------MTMTASGLQSLPQTTVGTAS---------
TC19278          ----------------------------------------------------------
TA14134          ----------------------------------GQAQNNDVQ----NSSSSVI---------
TC68930          ----------------------------------GNHSAVSQL----QSGAGGG---------
TC140470         ----------------------------------SEYPIPISHSADVQHKANSG---------
TC253044         ----------------------------------SEYPVPISHSADAQNKANSA---------
Os06g09370       ----------------------------------SEY--PINHSSDVEPQANSA---------
```

FIGURA 21 CONT.

```
AT2G24260          ---------AAPPQSRTKIRARRGQATDPHSIAERLRRERIAERMKALQELVPNGN---KT
AT4G30980          ---------AQP-QTKPKVRARRGQATDPHSIAERLRRERIAERMKSLQELVPNGN---KT
AT5G58010          ---------APVVRQKPRVRARRGQATDPHSIAERLRRERIAERMKSLQELVPNTN---KT
TA12416            ---------NTPAQQPKQRVRARRGQATDPHSIAERLRRERIAERMKALQELVPNAN---KT
TA5414             ---------NTPAQQPKQRVRARRGQATDPHSIAERLRRERIAERMKALQELVPNAN---KT
TC205173           ----------PAQPRQRVRARRGQATDPHSIAERLRRERIAERMKALQELVPNAN---KT
GSVIVT00017237001  ----------------RVRARRGQATDPHSIAERLRRERIAERMKALQELVPNAN---KT
TC7102             ---------TPAAQPKQQRVRARRGQATDPHSIAERLRRERIAERMKALQELVPNAN---KT
TA34455            ---------TPAAQPKQQRVRARRGQATDPHSIAERLRRERIAERMKALQELVPNAN---KT
TC7103             ----------QPKQQKVRARRGQATDPHSIAERLRRERIAERMKSLQELVPNAN---KT
TA37666            ----------QPKQQKVRARRGQATDPHSIAERLRRERIAERMKSLQELVPNAN---KT
TC67603            ----------KPKQRVRARRGQATDPHSIAERLRRVRIAERMKALQELVPNAN---KT
Os06g08500         ----------------------------LRRERIAERMKSLQELVPNAN---KT
Os02g55250         ----------GGAAAPPRQRQRARRGQATDPHSIAERLRRERIAERMKALQELVPNAN---KT
Os09g25040         ----------GGAPAPPRQRVRARRGQATDPHSIAERLRRERIAERMKALQELVPNANKLMQT
TA2544             ----------GAGTAPPKPRVRARRGQATDPHSIAERLRRERIAERMKALQELVPNAN---KT
DV982110           ----------AAPPVRPRVRARRGQATDPHSIAERLRRERIAERMKALQELVPNSN---KT
WO051050seqID516   ----------AAPPVRPRVRARRGQATDPHSIAERLRRERIAERMKALQELVPNSN---KT
AT4G02590          ----------QPTSIRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPTVN---KT
TC10015_part       ----------QPTSIRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPTVN---KT
AT1G03040          ----------QQSTIRPRVRARRGQATDPHSIAERLRRERIAERIRSLQELVPTVN---KT
TA13791            ----------PP-AMRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVN---KT
TC60118            ----------PP-AMHPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVN---KT
TC60119            ----------PP-TMRPRVRARRGQATDPHSIAERLRRERITERIRALQELVPSVN---KT
TC61833            ----------PP-TMRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVN---KT
TA3392             ----------PP-AMRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVN---KT
WO051050seqID1671  ----------PP-AMRPRVRARRGQATDPHSIAERLRRERIAERIRQLQDLVPSVN---KT
Pt_TC63334         ----------PP-AMRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVT---KT
TA3263             ----------PP-AMRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVN---KT
Pt_scaff_II.416    ----------PP-AMRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVN---KT
TA2825             ----------PP-TMRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVN---KT
TC207545           ----------PP-AVRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVN---KT
TC229602           ----------PP-AIRPRVRAIRGQATDPHSIAERLRRERIAERIRALQELVPSVN---KT
TA46156            ----------PP-AIRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVN---KT
AK247217           ----------PP-AIRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVN---KT
```

FIGURA 21 CONT.

```
TA30646                  ----------PP-AIRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVN---KT
TC8633                   ----------PP-AIRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPNVN---KT
TC15501                  ----------PP-AIRPRVRARRGQATDPHSIAERLRRERIAERMKALQELVPSSN---KT
TA1140                   ----------PP-TIRPRVRARRGQATDPHSIAERLRRERIAERMRALQELVPSCN---KT
TA3490_part              ----------PP-SIRPRVRARRGQATDPHSIAERLRRERIAERMRALQELVPSCN---KT
CK293938                 ----------SP-AIRPRVRARRGQATDPHSIAERLRRERISERIKALQELVPSCN---KT
TC12771                  ----------PP-AIRPRVRARRGQATDPHSIAERLRRERISERIKALQELVPSCN---KT
TC172581                 ----------PP-SIRPRVRARRGQATDPHSIAERLRRERISERIKALQELVPSCN---KT
TA28621                  ----------PP-SIRPRVRARRGQATDPHSIAERLRRERISERIKALQELVPSCN---KT
TA21665                  ----------QP-AIRPRVRAQRGQATDPHSIAERLRRERISERIKALQELVPSCN---KT
TA46194                  ----------PP-AIRPKVRARRGQATDPHSIAERLRRERIAERMKALQELVPSSN---KT
DY268946                 ----------PP-AVRPRVRARRGQATDPHSIAERLRRERIADRMRALQELVPSCN---KT
CO123623                 ----------PP-AVRPRVRARRGQATDPHSIAERLRRERIAERMKALQELVPSCN---KT
Pt_scaff_70.65           ----------PP-GIRPRVRARRGQATDPHSIAERLRRVRITERVKALQELVPTCN---KT
Pt_scaff_XIII.403        ----------PP-GIRPRVRARRGQATDPHSIAERLRRVRITERVKALQELVPTCN---KT
TA1616                   ----------PP-AIRPRVRARRGQATDPHSIAERLRRERIAERMKALQELVPTAN---KT
TA55042                  ----------PP-GIRPRVRARRGQATDPHSIAERLRRERIAERMKALQELVPSIN---KT
WO051050seqID77          ----------QP-AIRPRARARRGQATDPHSIAERLRRERINERMKALQELVPSCN---KT
TC3698                   ----------PT-GPRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSTN---KT
Os03g58330               ----------QPPAPRPKVRARRGQATDPHSIAERLRRERIAERMRALQELVPNTN---KT
TA139285                 ----------QPPAPRPKVRARRGQATDPHSIAERLRRERIAERMRALQELVPNTN---KT
TabHLH11                 ----------QPPAPRPKVRARRGQATDPHSIAERLRRERIAERMRALQELVPNTN---KT
TC104646                 ----------QPPAPRPKVRARRGQATDPHSIAERLRRERIAERMRALQELVPNTN---KT
TA126400                 ----------QPPAPRPKVRARRGQATDPHSIAERLRRERIAERMRALQELVPNTN---KT
Os07g08440               ----------QQPPAPRPKVRARRGQATDPHSIAERLRRERIAERMRALQDLVPNTN---KT
Os02g35660               PKGNGSGSAGNGAPKPRVRARRGQATDPHSIAERLRREKISDRMKDLQELVPNSN---KT
DT543504                 ----------NGSDRTVKPRVRARRGQATDPHSIAERLRREKIAERMKNLQELVPNSN---KT
Pt_scaff28.86            ----------EGCNGTGKTRVRARRGHATDPHSIAERLRREKIAERMKNLQELVPNSN---KV
GSVIVT00016367001        ----------GGCNGTGKPRVRARRGQATDPHSIAERLRREKIAERMKNLQELVPNSN---KT
TC19278                  ------------------------------------MKNLQELVPNSN---KT
TA14134                  ----------AGENTPYLLKELLVCSQATDPHSIAERERRERIAKNLKSLQELVPNAN---KT
TC68930                  ----------NAANGAARPRVRARRGQATDPHSIAERLRRERIAERMKSLQELVPNSN---KT
TC140470                 ----------NGNSASAKPRARARRGQATDPHSIAERLRREKISERMKNLQDLVPNSN---KA
TC253044                 ----------NGNSASAKPRARARRGQATDPHSIAERLRREKISERMKNLQDLVPNSN---KA
Os06g09370               ----------PGNSANAKPRTRARRGQATDPHSIAERLRREKISERMKNLQVLVPNSN---KA
                                                                  :: ** ***. .    :.
```

FIGURA 21 CONT.

```
AT2G24260                DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAASVSSQIS----EAGGSH-
AT4G30980                DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAASASSQIS----EDAG---
AT5G58010                DKASMLDEIIEYVRFLQLQVK----------VLSMSRLGGAGSVGPRLNGLSAEAGGRLN
TA12416                  DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAAAVAPLVADMSSEG-GGGD
TA5414                   DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAAAVAPLVADMSSEGAGGGD
TC205173                 DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAAAVAPLVADMSSEG--GGD
GSVIVT00017237001        DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAAAVAPLVADMSSEG--GGD
TC7102                   DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAAAVAPLVADMSSEG-RG--
TA34455                  DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAAAVAPLVADMSSEG-RG--
TC7103                   DKASMLDEIIDYVRFLQLQVK----------VLSMSRLGGAAAVAPLVADRSSEG-GGGD
TA37666                  DKASMLDEIIDYVRFLQLQVK----------VLSMSRLGGAAAVAPLVADRSSEG-GG-D
TC67603                  DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAAAVAPLVS----EG--GGD
Os06g08500               DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAGMAPLVASMSSE--GNSN
Os02g55250               DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGASAVAPLVANMSSESNGNGN
Os09g25040               DKASMLDEIIDYVKFLQLQVKASTYTKLLIHVLSMSRLGGAAAVAPLVADMSSE--GRGG
TA2544                   DKASMLDEIIDYLKFLQLQVK----------VLSMSRLGGAAAVAPLVADMSSE------
DV982110                 DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAGAVAPLVADIPAEG-----
WO051050seqID516         DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAGAVAPLVADIPAEG-----
AT4G02590                DRAAMIDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAPLVT-DMPLSSSVED
TC10015_part             DRAAMIDEIVDYVKFLRLQVK----------VLSMSRLGGAGAGAPLVT-DMPLSSSVED
AT1G03040                DRAAMIDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAPLVT-EMPLSSSVED
TA13791                  DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGAAGAVAPLVT-DLPLSS-VED
TC60118                  DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGAAGAVAPLVT-DLPLSS-VED
TC60119                  DRAVMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAPLVT-DIPLSS-VEY
TC61833                  DRAVMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAPLVT-DIPLSS-VED
TA3392                   DRAAMLDEIVDYVKFLRLQVK----------VLSMSRVGAPGAVAPLVTTDLPLSS-VED
WO051050seqID1671        DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGAAGAVAPLVT-DIPLSS-VEE
Pt_TC63334               DRATMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAPLVT-DIPLSS-VED
TA3263                   DRATMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAPLVT-DIPLSS-VED
Pt_scaff_II.416          DRAAMLDEIVDYVKFLRLQVK----------ILSMSRLGGAGAVAPLVT-DIPLSP-VED
TA2825                   DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAPLVT-DIPLSS-VED
TC207545                 DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAPLVT-DIPLSS-VEE
TC229602                 DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAPLVT-DIPLSS-VEE
TA46156                  DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAPLVT-DIPLAS-VEE
AK247217                 DRAVMLDEIVDYIKFLRLQVK----------VLSMSRLGGAGAVAPLVT-DIPISS-VEE
```

FIGURA 21 CONT.

```
TA30646                 DRAVMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAPLVT-DIPISS-VEE
TC8633                  DRAIMLDEIVDYVKFLRLQVK----------VLSMSRLGGAAAVAPLVT-EIPISS-VEE
TC15501                 DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAQLVA-DIPLSS-AEG
TA1140                  DKAAMLDEILDYVKFLRLQVK----------VLSMSRLGGAGAVAQLVS-DVPLQS-VEG
TA3490_part             DKAAMLDEILDYVKFLRLQVK----------VLSMSRLGGAGAVAQLVS-DVPLQS-VEG
CK293938                DRAAMLDEILDYVKFLRLQVK----------VLSMSRLGGASAVAQLVA-DIPLQS-MEG
TC12771                 DRAAMLDEILDYVKFLRLQVK----------VLSMSRLGGASAVAQLVA-DIPLQS-VEG
TC172581                DRAAMLDEILDYVKFLRLQVK----------VLSMSRLGGASAVAQLVA-DIPLNL-WEG
TA28621                 DRAAMLDEILDYVKFLRLQVK----------VLSMSRLGGASAVAQLVA-DIPLQS-VEG
TA21665                 DRAAMLDEILDYVKFLRLQVK----------VLSMSRLGGTSAAAQVVA-DIPLQS-VEG
TA46194                 DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAQLVA-DIPLPA-VEG
DY268946                DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGAAGAVAQLVA-DVPLSSALE-
CO123623                DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGAAGAVAQLVA-DVPLLS-VE-
Pt_scaff_70.65          DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGAAGAVAQLVA-DVPLSSVQ--
Pt_scaff_XIII.403       DRAAMLDEIVDYVKFLRLQIK----------VLSMSRLGAAGAVAQLVA-DVPLSSVQIK
TA1616                  DRAAMIDEIVDYVKFLRLQVK----------VLSMSRLGAAGAVAQLVA-DVPLASVE--
TA55042                 DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAQLVA-DVPLSAVEG-
WO051050seqID77         DRAAMLDEIVDYVKFLRLQVK----------VLSMSRLGGAGAVAQLVA-DVPLSSAEG-
TC3698                  DRAVMLDEIVEYVKFLRLQVK----------VLSMSRLGGAGAVAQLVA-DIPMSSSVEG
Os03g58330              DRAAMLDEILDYVKFLRLQVK----------VLSMSRLGGAGAVAQLVA-DIPLSV--KG
TA139285                DRAAMLDEILDYVKFLRLQVK----------VLSMSRLGGAGAVAQLVA-DIPLSV--KG
TabHLH11                DRAVMLDEILDYVKFLRLQVK----------VLSMSRLGGAGAVAQLVS-DIPLSV--KG
TC104646                DRAAMLDEILDYVKFLRLQVK----------VLSMSRLGGAGAVAQLVA-DIPLSV--KG
TA126400                DRAAMLDEILDYVKFLRLQVK----------VLSMSRLGGAGAVAQLVA-DIPLSV--KG
Os07g08440              DRAAMLDEILDYVKFLRLQVK----------VLSMSRLGGAGAVAQLVA-DIPISV--KG
Os02g35660              NKASMLDEIIDYVKFLQLQVK----------VLSMSRLGAAEAVVPLITETQTESPGFLL
DT543504                DKASILDEIIGYVKFLQLQVK----------VLSMSRLGAAAAVVPLITDGRAEVSN--G
Pt_scaff28.86           DKASMLDEIIEYVKFLQLQVK----------VLSMSRLGAAGAVIPLITDGQPEGHN--S
GSVIVT00016367001       DKASMLDEIIEYVKFLQLQVK----------VLSMSRLGAAEAVVPLITDGQAEGSK--G
TC19278                 DKASMLDEIIEYVRFLQLQVK----------VLSMSRLGATGAVVPLITENQPEGST--S
TA14134                 DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGGAGAVAPLIADVPAEGSG--S
TC68930                 DKASMLDEIIEYVKFLQLQVK----------VLSMSRLGGAGAVAPLIADVPSQGSG--G
TC140470                DKSSMLDEIIDYVKFLQLQVK----------VLSMSRLGAPGAVLPLLAESQTEGRSNSP
TC253044                DKSSMLDEIIDYVKFLQLQVK----------VLSMSRLGAPGAVLPLLAESQTEGRSNSP
Os06g09370              DKASMLDEIIDYVKFLQLQVK----------VLSMSRLGAPGAVLPLLRESQTECHSNPS
                        ::: ::***: *::**:**:*          :*****:*.. .    :
```

FIGURA 21 CONT.

```
AT2G24260          --GNASSAMVGGS----------QTAGNSNDSVTMTEHQVAKLMEEDMGSAMQYLQGK
AT4G30980          ----------GS----------HENTSSSGEAKMTEHQVAKLMEEDMGSAMQYLQGK
AT5G58010          ALTAPCNGLNGNG----------NATGSSNESLRSTEQRVAKLMEEDMGSAMQYLQGK
TA12416            CIQANGR-NPNG----------AQTTSANDSLTVTEHQVAKLMEEDMGSAMQYLQGK
TA5414             CIQANGR-NPNG----------AQTTSANDSLTVTEHQVAKLMEEDMGSAMQYLQGK
TC205173           CIQANGKSNGGAQASTTNTNTNQTTATTTSNDSLTMTEHQVAKLMEEDMGSAMQYLQGK
GSVIVT00017237001  CIQASGTSGPTGGRATNG------TQTTTSNDSLTVTEHQVAKLMEEDMGSAMQYLQGK
TC7102             --EGNGGRGG-NG----------TASSSNN-DSMTVTEHQVAKLMEEDMGSAMQYLQGK
TA34455            --EGNVGRGG-NG----------TAASSSNKETMTVTEHQVAKLMEEDMGSAMQYLQGK
TC7103             CAQANGGGRGSNG----------TTSLANNDS-MTMTEHQVAKLMEEDMGSAMQYLQGK
TA37666            CVQGNGGRGGSNG----------TTSSANNDSSMTMTEHQVAKLMEEDMGSAMQYLQGK
TC67603            CIQTSANGGSHPRNSN--------GDQTPSANDRLTMTEHQVAKLMEEDMGSAMQYLQEK
Os06g08500         GSSNGGGGKASKGGTGGEGGGGGGGGGGTGGGMRVTEQQVAKMMEEDMGTAMQYLQGK
Os02g55250         ATSSSGNGEAANGSSNGDNNGGG----------TLRVTEQQVAKLMEEDMGSAMQYLQGK
Os09g25040         GAANGGAPAAAGSDS----------LTVTEQQVAKLMEEDMGTAMQYLQGK
TA2544             --SGGGAGTPRSNNDG----------LTAAEQQVAKLLEEDMGSAMQYLQGK
DV982110           --ANAPQGTRTNG----------SQNSSPDGLALTERQVAKLMEEDMGTAMQYLQGK
WO051050seqID516   --ANAPQGTRTNS----------SQNSSPDGLALTERQVAKLMEEDMGSAMQYLQGK
AT4G02590          ETGEGGRTP----------QPAWEKWSNDGTERQVAKLMEENVGAAMQLLQSK
TC10015_part       ESGEGGRAP----------QPAWEKWSNDGTERQVAKLMEENVGAAMQLLSSK
AT1G03040          ET----------QAVWEKWSNDGTERQVAKLMEENVGAAMQLLQSK
TA13791            ESGEGGRS----------QPAWEKWSNDGTERQVAKLMEEDVGAAMQFLQSK
TC60118            ESGEGGRS----------QPAWEKWSNDGTERQVAKLMEEDVGAAMQFLQSK
TC60119            ESGEGGRS----------QPAWEKWSNDGTEQQVAKLMEENVGAAMQFLQSK
TC61833            ESGDGGRN----------QPAWEKWSNDGTERQVAKLMEENVGAAMQFLKSK
TA3392             ESGEGVRN----------QPAWEKWSNDGTERQVAKLMEENVGAAMQFLQSK
WO051050seqID1671  EGGEGGRN----------QPAWEKWSNDGTERQVAKLMEENVGAAMQFLQSK
Pt_TC63334         ETGEGGRN----------QPAWEKWSNDGTERQVAKLMEENVGAAMQFLQSK
TA3263             ETGEGGRN----------QPAWEKWSNDGTERQVAKLMEENVGAAMQFLQSK
Pt_scaff_II.416    ETGEGGRN----------QLAWEKWSNDGTERQVAKLMEENVGAAMQFLQSK
TA2825             ETGEGGRN----------QPAWEKWSNDGTERQVAKLMEENVGAAMQFLQSK
TC207545           EGGEG-RN----------QPAWEKWSNDGTERQVAKLMEENVGAAMQFLQSK
TC229602           EGGEGARN----------RPAWDKWSNDGTERQVAKLMEENVGAAMQFLQSK
TA46156            EASEGGRN----------EPAWEKWSNDGTERQVAKLMEENVGAAMQFLQSK
AK247217           ESSEGGNNN----------QPAWEKWSSDGTERQVAKLMEENVGAAMQFLQSK
```

FIGURA 21 CONT.

```
TA30646                 ESSEGGNNN----------------------QPAWEKWSSDGTERQVAKLMEENVGAAMQFLQSK
TC8633                  EISEGGNN-----------------------QPAWEKWSSDGTERQVAKLMEENVGSAMQFLQSK
TC15501                 GASEGGSN-----------------------QPAWEKWSTDGTEHQVAKLMEEDVGAAMQFLQSK
TA1140                  DGSENGYNNQQ-Q------------------GQAAWENWSNDDTEREVAKLMEEDVGAAMQFLQSK
TA3490_part             DTSENGYN-----------------------QPAWENWSNDDTEREVAKLMEEDVGAAMQFLQSK
CK293938                DSAESKSN-----------------------QHIWEKWSNVDTEQEVAKLMEEDVGAAMQYLQSK
TC12771                 DSAESRSN-----------------------QHIWEKWSNVDTEQEVAKLMEEDVGAAMQYLQSK
TC172581                DSGESRSN-----------------------QHIWDKWSNVDTEREVAKLMEEDVGAAMQYLQSK
TA28621                 DSGESRSN-----------------------QHIWDKWSNVDTEQEVAKLMEEDVGAAMQYLQSK
TA21665                 DTCESHSN-----------------------QHVWEKWSDSETEQEVAKLMEEDVGTAMQYLQSK
TA46194                 ETGEGGSN-----------------------QQAWDKWSNDGTEREVAKLMEEDVGAAMQFLQSK
DY268946                GESIDGG-----S------------------SQPEWEKWSNDGTEQQVAKLMEEDIGAAMQFLQSK
CO123623                GDGAGGG------------------------TQPPWEKWSNDGTEQQVAKLMEEDIGAAMQFLQSQ
Pt_scaff_70.65          GEGIEGG-----A------------------NQQAWENWSNDGTEQEVAKLMEEDVGAAMQLLQSK
Pt_scaff_XIII.403       GEGNEGG-----A------------------NQQSWENWSNDDTEQEVAKLMEEDVGAAMQFLQSK
TA1616                  GESIDGAA----A------------------NQQTWEKWSNDGTEQQVAKLMEEDIGAAMQFLQSK
TA55042                 DQDIEGG-----A------------------NEQAWDKWSNDGTEQQVAKLMEEDVGAAMQFLQSK
WO051050seqID77         -EIIEGG-----N------------------NQPAWEKWLTDGTEQQVAKLMEEDVGAAMQFLQSK
TC3698                  DSSESGKTS----------------------YQGWEKWSTDGTERQVAKLMEEDVGAAMQFLQSK
Os03g58330              EASDSG-------G------------------NQQIWEKWSTDGTERQVAKLMEEDIGAAMQFLQSK
TA139285                EASDSG-------S------------------KQQIWEKWSTDGTERQVAKLMEEDIGAAMQFLQSK
TabHLH11                EASDGG-------S------------------KQQIWEKWSTDGTEKQVAKLMDEDIGAAMQFLQSK
TC104646                EASDSG-------S------------------TQHIWEKWSTDGTEKQVAKLMEEDIGAAMQFLQSK
TA126400                EAGDGGGAPQQQQ------------------QQHVWEKWSTDGTEKQVAKLMEEDIGAAMQFLQSK
Os07g08440              EASDSG-------S------------------KQQIWEKWSTDGTEKQVAKLMEEDIGAAMQFLQSK
Os02g35660              SPRSSSGERQAGAGAVTG----GLPGDQPELLDGGAMFEQEVVKLMEDNMTTAMQYLQSK
DT543504                LSLAPLAG-----------------------QGVDFSPSPDQVVFEQEVVKLMESNMTMAMQYLQSK
Pt_scaff28.86           LSLSPSAG-----------------------LGIDISPSADQIAFEQEVLKLLESDVTMAMQYLQSK
GSVIVT00016367001       LSLSPSAG-----------------------QAEDICQSPDQIAFEQEVVKLMESNVTMAMQYLQSK
TC19278                 HFLSQSAG-----------------------RDVKHIESPNTLTFEQEVVKLMESNMTKAMQFLQRK
TA14134                 LASAALGQ-----------------------AGG--SLSQDGLAFEQNVVKLMEKDMTAAMQYLQNK
TC68930                 VMSTALGQ-----------------------ASGPLTLSQDGLAFEQEVARLMESNMTSAMQYLQNK
TC140470                LSSPTTSQ-----------------------GLLDVAGPEDSLVFEQEVIKLMETSITNAMQYLQNK
TC253044                LSSPTASQ-----------------------GLLDAAGPEDSLVFEQEVIKLMETSITNAMQYLQNK
Os06g09370              LSASTISQ-----------------------GPPDMPDSEDSSAFEQEVVKLMETSIISAMQYLQNK
                                                        *:.* :::: .:  *** *. :
```

FIGURA 21 CONT.

```
AT2G24260                GLCLMPISLATAISTATCHSRNPLIPG----------------------------------AVADVG
AT4G30980                GLCLMPISLATTISTATCPSRSPFVK----------------------------------DTG
AT5G58010                GLCLMPISLATAISSSTTHSRGSLFN----------------------------PISSAV
TA12416                  GLCLMPISLATAISTATCHSRNPIISTSNNNNNNN----GNPH---------HNPLLQSNGE
TA5414                   GLCLMPISLATAISSATCHSRNPIISTSNNNNNNN----GNPH---------HNPLFQSNGE
TC205173                 GLCLMPISLATAISTATCPTRNVNVNPLINAAA-----GATH---------FPTAANPAGE
GSVIVT00017237001        GLCLMPISLATAISTTTCHSRNPMVAAAAVAASNI-NNGSHT---------HPLLPNSNAD
TC7102                   GLCLMPISLATAISTATCHSRNPMIPNGNGNNPLLGGGGGL---------ATNGGGGEAG
TA34455                  GLCLMPISLATAISTS---TRIPLLSP----------------------------EAG
TC7103                   GLCLMPISLATAISTATCHSMKLNNPLLHGGGSG-----GSA---------AINGGG---G
TA37666                  GLSLMPISLATAISTSTCHSMKPNNPLILGGGS----------------AINGGGETGG
TC67603                  GLCLMPISLATAISSATSRSRNPMINHENPANGS----------------HLLIQSSGGD
Os06g08500               GLCLMPISLASAISSATSSA-------SLLSRPSI----------------RHAGAPPQTM
Os02g55250               GLCLMPISLATAISSATSS-------SLLPRTGG----------------GAGGSLHEGG
Os09g25040               GLCLMPISLASAISSATCHLRPPVVAAAQQFPAG----------------LGAAAAAAHH
TA2544                   GLCLMPISLASAISSATCRP---------RPPPG----------------VNLRYPAAGD
DV982110                 GLCLMPISLASAINNSGSRPQAPSTPSXXXLLAS----------------NVNDRQVL
WO051050seqID516         GLCLMPISLASAINNSGARPQAPSTPSLQGLLPP----------------NANDRQIL
AT4G02590                ALCMMPISLAMAIYHSQPPDTSS----------------------------------VV
TC10015_part             ALCMMPISLAMAIYHSQPPDTT-----------------------------------IV
AT1G03040                ALCIMPISLAMAIYHSQPPDTSSS---------------------------------IV
TA13791                  ALCVMPISLATAIYHTQSPDTSS----------------------------------VV
TC60118                  ALCVMPISLATAIYHTQSPDTSS----------------------------------VV
TC60119                  SLCIMPISLATAIYHTQVPDTSS----------------------------------VV
TC61833                  ALCIMPISLATAIYHSQPLDTSS----------------------------------IV
TA3392                   ALCIMPISLATAIYHSQPPE-------------------------------------
WO051050seqID1671        ALCIMPISLATAIYQTQPPDTSS----------------------------------LV
Pt_TC63334               ALCIMPISLATAIYHTQPPDTTT----------------------------------IV
TA3263                   ALCIMPISLATAIYHTQPPDTTS----------------------------------IV
Pt_scaff_II.416          ALCIMPITLATAIYHTQPPDTTT----------------------------------IV
TA2825                   ALCIMPISLATAIYHTQAPDTST----------------------------------IV
TC207545                 ALCIMPVSLASAIYQSQPSGTSS----------------------------------IV
TC229602                 ALCIMPISLASAIYQSQPPDTSS----------------------------------IV
TA46156                  ALCIMPISLASAIYHSQQPDTTP----------------------------------LI
AK247217                 ALCIMPISLASAIYHSQPPDTSS----------------------------------LV
```

FIGURA 21 CONT.

```
TA30646                 ALCIMPISLASAIYHSQPPDTSS----------------------------------------LV
TC8633                  ALCIMPISLASAIYHSQPPDTSS----------------------------------------LI
TC15501                 ALCIMPISLASAIYHDQPPDAST----------------------------------------MI
TA1140                  ALCIMPISLASLIYPNHQPDVKP----------------------------------------EP
TA3490_part             ALCIMPISLASLIYPPQTPDSNS----------------------------------------HV
CK293938                SLCIMPISLAALIYPTQQPDDMS----------------------------------------MV
TC12771                 SLCIMPISLAALIYPTQQPDDPS----------------------------------------MV
TC172581                SLCIMPISLAALIYPTQQPDDQS----------------------------------------LV
TA28621                 SLCIMPISLAALIYPTQQPDDQS----------------------------------------LV
TA21665                 SLCIMPISLAALIYPTQQSDNQS----------------------------------------MV
TA46194                 ALCIMPISLAAAIYPAHQTDTPT----------------------------------------LI
DY268946                ALCIMPISLASAIYRTRQPDAPA----------------------------------------FV
CO123623                ALCIVPISLASAIFPAHQPDAPT----------------------------------------IV
Pt_scaff_70.65          ALCIMPVSLASAIFRARPPNAPT----------------------------------------LV
Pt_scaff_XIII.403       ALCIMPISLASAIFRARPPNAST----------------------------------------LI
TA1616                  ALCIMPISLASAILRTHPPDAPS----------------------------------------II
TA55042                 ALCIMPISLASAIFRMPQSEAST----------------------------------------GI
WO051050seqID77         TLCIMPISLASAIFRTSQPDMPR----------------------------------------SI
TC3698                  ALCIMPISLAAAICQTNPSTEISF----------------------------------------M-
Os03g58330              ALCMMPISLAMAIYDTQQTQDGQ----------------------------------------PV
TA139285                ALCMMPISLAMAIYDTQHSQDGQ----------------------------------------PV
TabHLH11                ALCMMPVSLAMAIYDTQHSQDGQ----------------------------------------PV
TC104646                ALCMMPISLAMAIYDTQHSQDGH-S--------------------------------------LM
TA126400                ALCMMPVSLAMAIYDTQHPLDGHGH--------------------------------------SL
Os07g08440              ALCMMPISLAMAIYDTQHSQDGH----------------------------------------SV
Os02g35660              GLCLMPVALASAISAQKGTSSAAVRPEKKKNGDG-----------------------DGGGDEEDV
DT543504                GFCLMPVALAAAISNVK-----------SSTS-------------------------------S
Pt_scaff28.86           GLCLMPIALAAAISSVK----------ASLSGTTSEERKNNGYT-------------------SGLVS
GSVIVT00016367001       GLCLMPIALATAISSGK----------AASSGTGSDEGK-----------------------N
TC19278                 GLCLMPIALAAAISGSK----------APPASLSTEGKKSILTNGFVHHNSGQSNIGPGQIS
TA14134                 GLCLMPIALATAISSSTGKPLLGSGVTNSVDSTSDRQSSGNQPASLTVDSCSGALGMGSA
TC68930                 GLCLMSIALATAISSSSGKPVLATVPGTGVDS-SEKQNSDMQSIVLPFSSSSTTMGLRAT
TC140470                GLCLMPIALASAISNQKG---------------------------------------------T
TC253044                GLCLMPIALASAISNQKG---------------------------------------------T
Os06g09370              GLCLMPIALASAISNQKGM--------------------------------------------A
                          :.::.::** *
```

FIGURA 21 CONT.

```
AT2G24260          GPSPPNLSGMTIQ--STSTKMG------SGNGKLNGNGVTERSSSIAVKEAVSVSKAATG
AT4G30980          VPLSPNLS----------TTIVA------NGNG----------SSLVTVKDAPSVSKP---
AT5G58010          AAEDSNVT----------ATAVA------APEAS----------STMDDVSASKA---
TA12416            GPTSPSMS----------VLTVQS------ATMGNG---------GADGSVKDAASVSKP---
TA5414             GPTSPSMS----------VLTVQS------ATMGNG---------GADGSVKDAASVSKP---
TC205173           GPSSPSMS----------VLTVQS------AIAGN----------DGAASVSKP---
GSVIVT00017237001  GPSSPSMS----------VLTVQS------ATMGNG---------LADAPVKDAASVSKP---
TC7102             GPSSPKLS----------VLTVQS------ATIGNG---------GIDPSVKDATSVFEA---
TA34455            RPASPTLS----------ALTVQS------ATIGN----------AAGKDATSLFEA---
TC7103             GPSSPSLS----------ASTVQS------ATMGN----------GGA---------
TA37666            GPSSPTLS----------ASTVQS------ATMGN----------GGT---------
TC67603            GPSSPSMS----------VLTVQS------ATMGNG---------GLEG---GAASVSKP---
Os06g08500         LDAAGPTS----------PAAMS------NGDDP----------RHAKADGGAGGTQ---
Os02g55250         NGTSPPLV----------NGTAT------GCDDAGVFFSVKFVVELSFLLLNEDCRGKEESK
Os09g25040         HQLSAAAA----------AAAMRGHLPGLNADGSVPASPSMSVLTAQSAMANGGGGAADGEG
TA2544             APPSPSIS----------ALTVQ----------SSNAGSAGADAS-
DV982110           EPS---------------------------------------------------------
WO051050seqID516   EPSNSALSALTSTSMTIQSSISSPGSGITEQGDNAHKAVNGTRNPKDSREANSVTKSNGI
AT4G02590          KPENNPPQ----------------------------------------------------
TC10015_part       ---------------------------------------------------------
AT1G03040          KPEMNPPP----------------------------------------------------
TA13791            KPETNPPS----------------------------------------------------
TC60118            KPETNPPS----------------------------------------------------
TC60119            KPETNPPA----------------------------------------------------
TC61833            KPETDPPP----------------------------------------------------
TA3392             ---------------------------------------------------------
WO051050seqID1671  KSESNPPS----------------------------------------------------
Pt_TC63334         KPETNPPS----------------------------------------------------
TA3263             KPETNPQS----------------------------------------------------
Pt_scaff_II.416    KPETNPPS----------------------------------------------------
TA2825             KPETNPPS----------------------------------------------------
TC207545           KPETNPPS----------------------------------------------------
TC229602           KPETNPPS----------------------------------------------------
TA46156            KPQTNPPS----------------------------------------------------
AK247217           KPETNPPS----------------------------------------------------
```

FIGURA 21 CONT.

```
TA30646                 KPETNPPS----------------------------------------------------
TC8633                  KPETNPPS----------------------------------------------------
TC15501                 KPEPNAPS----------------------------------------------------
TA1140                  SAPS--------------------------------------------------------
TA3490_part             KPETV-------------------------------------------------------
CK293938                KPEPAAPS----------------------------------------------------
TC12771                 KPEPAAPS----------------------------------------------------
TC172581                KPEAAAPS----------------------------------------------------
TA28621                 KPEAAAPS----------------------------------------------------
TA21665                 KPEQAAPL----------------------------------------------------
TA46194                 KPEPNAPS----------------------------------------------------
DY268946                KPESSTPSYCLNNNWSQQLAHNDSKYLITGTYSYRILLVLLSNQIIQIPMQRSHSSFP--
CO123623                ----------------------------------------------------------
Pt_scaff_70.65          KPESNPPS----------------------------------------------------
Pt_scaff_XIII.403       NTESNTPS----------------------------------------------------
TA1616                  KPESNTPS----------------------------------------------------
TA55042                 KPESNSH-----------------------------------------------------
WO051050seqID77         KPESSAPS----------------------------------------------------
TC3698                  ----------------------------------------------------------
Os03g58330              KHEPNTPS----------------------------------------------------
TA139285                KPEPNTPS----------------------------------------------------
TabHLH11                KPEPNNTA----------------------------------------------------
TC104646                KPEPNTSS----------------------------------------------------
TA126400                KPEPNASS----------------------------------------------------
Os07g08440              KPEPNTPS----------------------------------------------------
Os02g35660              KGEFDAPRRPPVGRPKEMRSRV---------------------------------------
DT543504                SSSSVPASD---------------------ESKKLGLH--QQYSNQQHLQQ----------
Pt_scaff28.86           SSSSITGIDTHPMSNDNNIATGTLSSKGMIVNGCNEVVKQEVLKNT----------------
GSVIVT00016367001       AATPVAATP---------------------VATAYLALG---HIIHLLMAMF--------
TC19278                 SDGETRSEENIIGIHSREDFSSNSCSGTSIKKEGTNTTNFTREMNQE--------------
TA14134                 GFGSDFLLKENTVEKRGRELVPTAESNGAEFGIISSLPKLRKKE----------------
TC68930                 ASGSDAPINENSINKASIEKVVAEKSNGISPGLSSDVPKVGSQSREELLKRAQ--------
TC140470                SAAAIPPER---------------------------------------------------
TC253044                SAAAIPPER---------------------------------------------------
Os06g09370              AAAAIPPEK---------------------------------------------------
```

FIGURA 21 CONT.

```
AT2G24260          MQFDESDARVWIWFEIRREDDIVELLWQSGQVVGTNQTHRQSYDPPPILRGSGSGRGEEN
AT4G30980          ------------------------------------------------------------
AT5G58010          ------------------------------------------------------------
TA12416            ------------------------------------------------------------
TA5414             ------------------------------------------------------------
TC205173           ------------------------------------------------------------
GSVIVT00017237001  ------------------------------------------------------------
TC7102             ------------------------------------------------------------
TA34455            ------------------------------------------------------------
TC7103             ------------------------------------------------------------
TA37666            ------------------------------------------------------------
TC67603            ------------------------------------------------------------
Os06g08500         ------------------------------------------------------------
Os02g55250         LLVQKGP-------------------------------------------------------
Os09g25040         SQLKDAASVSKP--------------------------------------------------
TA2544             ------------------------------------------------------------
DV982110           ------------------------------------------------------------
WO051050seqID516   GGPALSKNAVKGEDEPQRRTGQ-----------------------------------------
AT4G02590          ------------------------------------------------------------
TC10015_part       ------------------------------------------------------------
AT1G03040          ------------------------------------------------------------
TA13791            ------------------------------------------------------------
TC60118            ------------------------------------------------------------
TC60119            ------------------------------------------------------------
TC61833            ------------------------------------------------------------
TA3392             ------------------------------------------------------------
WO051050seqID1671  ------------------------------------------------------------
Pt_TC63334         ------------------------------------------------------------
TA3263             ------------------------------------------------------------
Pt_scaff_II.416    ------------------------------------------------------------
TA2825             ------------------------------------------------------------
TC207545           ------------------------------------------------------------
TC229602           ------------------------------------------------------------
TA46156            ------------------------------------------------------------
AK247217           ------------------------------------------------------------
```

FIGURA 21 CONT.

FIGURA 21 CONT.

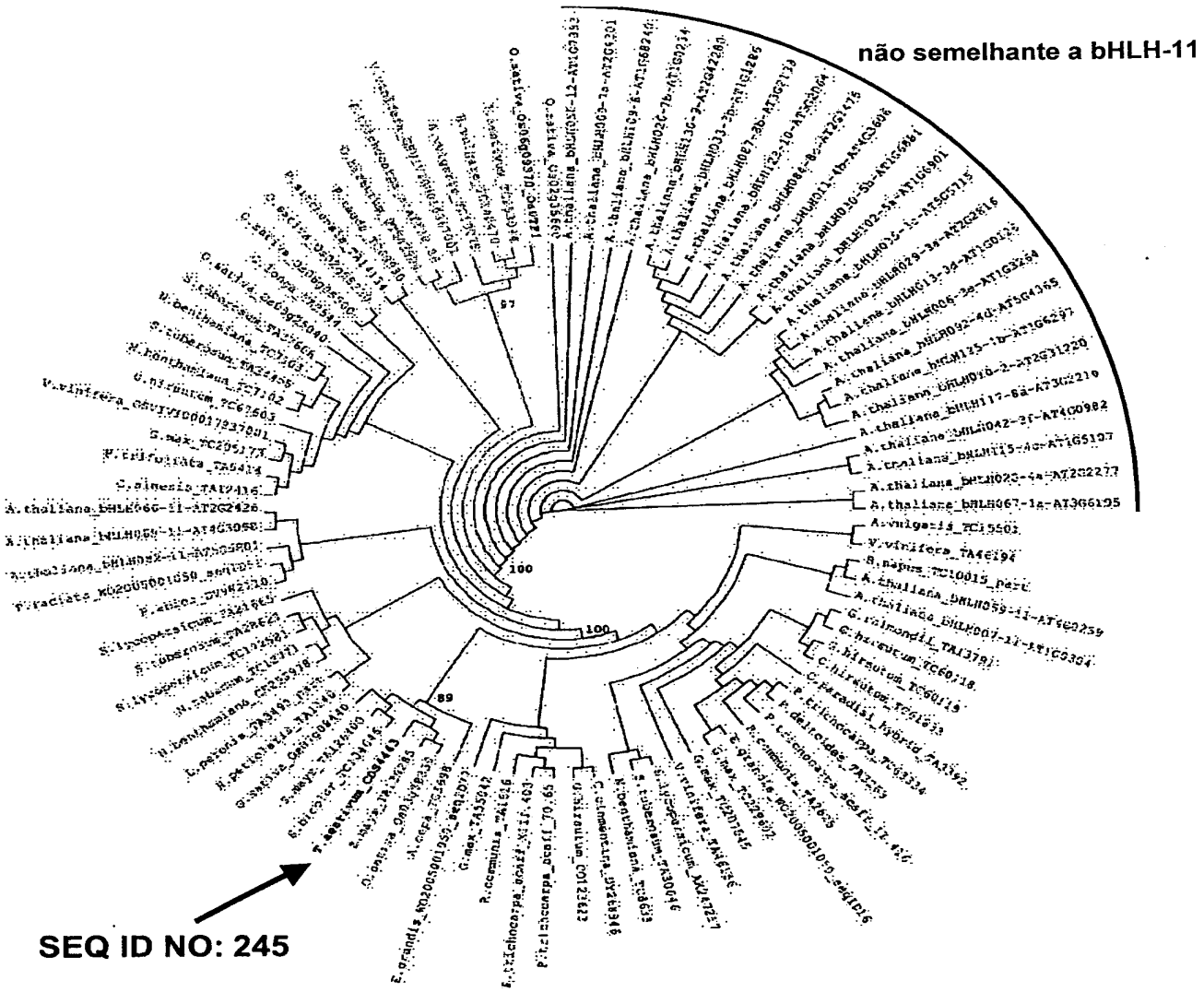


FIGURA 22

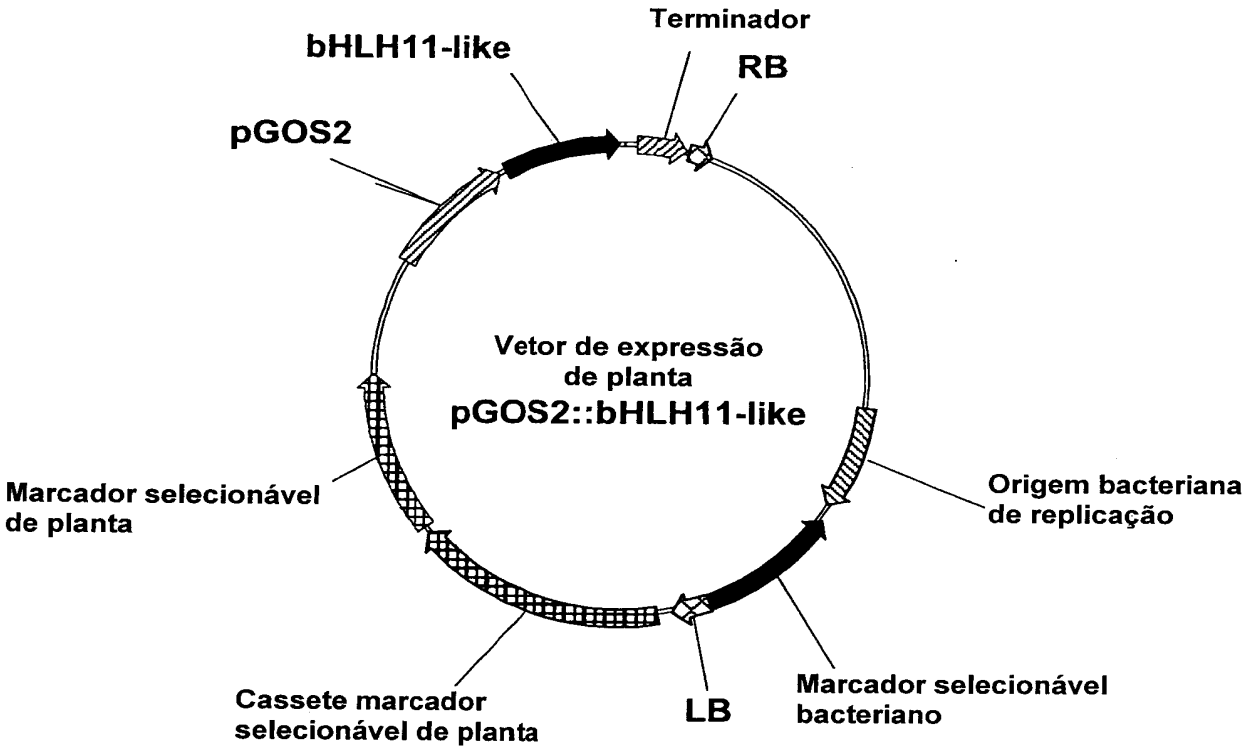

Terminador
bHLH11-like
RB
pGOS2
Vetor de expressão
de planta
pGOS2::bHLH11-like
Marcador selecionável
de planta
Origem bacteriana
de replicação
Cassete marcador
selecionável de planta
LB
Marcador selecionável
bacteriano

FIGURA 23

SEQ ID NO: 244, sequência codificadora de TabHLH11, Triticum aestivum

ATGGCGGGGCAGCCGCCGCAGGGCCCGGAGGACGACTTCCTCGACCAGTTCTTCTCCCTCACCAAC
TCCCTCTCCGCCGCCGGCCGCCCCTCCGGCGACCAGCCCTTCTCCCTCGCCCTCAGCCTCGACGCC
GCCTCGGACGCCTCCGGCAGCAGGGGCGGCATCGGCGACGACGCCGGCAACGCCGCGGAGCGGGAC
GGCGTGCAGCTCCCCGGCCTCTTCCCGCCGGTGTTCGGCGGTGGCCTCCAGCCGCCGCACCTTCGC
CCCAGCCACCCTCCCCCACAGATGTTCCACGCGCAGCAGCCGAAGCAGGGCGGGCCGGCTGGAGGG
CCGCAGCCGCCGGCGCCGAGGCCGAAGGTGCGGGCGCGTCGCGGTCAGGCGACTGATCCCCACAGC
ATCGCGGAGAGGCTAAGAAGAGAGAGGATAGCGGAAAGGATGAGGGCCCTACAGGAATTGGTCCCC
AATACGAACAAGACAGATAGGGCAGTTATGCTAGATGAGATCCTGGATTATGTGAAGTTCCTTAGG
CTTCAAGTAAAGGTGTTAAGCATGAGCAGATTGGGTGGTGCTGGTGCTGTTGCACAGCTGGTTTCT
GACATTCCACTTTCAGTTAAGGGGGAAGCAAGCGATGGTGGGAGCAAACAGCAGATATGGGAAAAG
TGGTCAACGGATGGCACGGAAAAACAGGTTGCGAAGCTGATGGACGAGGACATCGGTGCCGCAATG
CAATTTCTCCAATCAAAGGCTCTCTGCATGATGCCAGTCTCCCTTGCCATGGCTATCTATGACACA
CAACATTCACAGGACGGCCAACCAGTGAAGCCTGAACCCAACAACACTGCCTAG

SEQ ID NO: 245, TabHLH11, Triticum aestivum

MAGQPPQGPEDDFLDQFFSLTNSLSAAGRPSGDQPFSLALSLDAASDASGSRGGIGDDAGNAAERD
GVQLPGLFPPVFGGGLQPPHLRPSHPPPQMFHAQQPKQGGPAGGPQPPAPRPKVRARRGQATDPHS
IAERLRRERIAERMRALQELVPNTNKTDRAVMLDEILDYVKFLRLQVKVLSMSRLGGAGAVAQLVS
DIPLSVKGEASDGGSKQQIWEKWSTDGTEKQVAKLMDEDIGAAMQFLQSKALCMMPVSLAMAIYDT
QHSQDGQPVKPEPNNTA

SEQ ID NO: 246, Motivo 1

(E/D) (D/S/E) (F/M) (L/F) (D/E/Q/L) (Q/H/E)

SEQ ID NO: 247, Motivo 2

RA(R/I/Q)RG(Q/H)ATDPHSIAER

SEQ ID NO: 248, Motivo 3

(M/I/V/L) (K/R) (A/S/Q/D/N)LQ(E/D/V)LVP

SEQ ID NO: 249, Motivo 4

(M/I) (L/I)DEI(I/V/L) (D/E/G)Y(V/L/I) (K/R)FL(Q/R)LQ(V/I)K

SEQ ID NO: 250, Motivo 5

(V/I)LSMSR(L/V)G

SEQ ID NO: 251, Motivo 6

V(A/V/L/I) (K/R) (L/M) (M/L) (E/D) (E/D/S/K/T) (D/N/S) (M/V/I) (G/T/I)XAMQ
(Y/L/F)L

SEQ ID NO: 252, Motivo 7

(M/V) (P/S) (I/V) (S/T/A)LA

SEQ ID NO: 253, Motivo 8

SIAERLRRERIAERMRALQELVPNTNKTDRAVMLDEILDYVKFLRLQVKVL

FIGURA 24

SEQ ID NO: 254, prm009718 (fwd):
GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGGCGGGGCANCCG

SEQ ID NO: 255, prm009719 (rev):
GGGGACCACTTTGTACAAGAAAGCTGGGTATGGGTGTTGCAGCTGCTGTT

SEQ ID NO: 256, Promotor Gos2 de arroz
AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAATAT
AAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCACCTACT
TTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGTAATTAAGT
GGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGTTAAATTATTCG
AGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTTCTAGCTGAACTCAATGGGT
AAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTCTGAACGTATTGGCAAAGATTTAAA
CATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACATCATTAAGGACATGTCTTA
CTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATTTTTATTATTTATCTTTTTTCG
ATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAAT
ACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTAGGTAGCAATATC
TGAATTCAAGCACTCCACCATCACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTT
ACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAAAAAAGAATT
TTGCTCGTGCGCGAGCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCCACA
GAACAACCCACAAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAG
GCTTTGCGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGCATCCTCCTTCTCCCATCTATAA
ATTCCTCCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAAG
GACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTCGATCCATATCTTCCGGTCGAGTTCTTGGT
CGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCTCCCTTCGGTTGTTCTTGGATTT
ATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCC
TGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTCGGTTTGATTAGTAGT
ATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTAGGGATCGGAATCTTGCGATT
TTGTGAGTACCTTTTGTTTGAGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAGTAC
GGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTC
CCTATTGAACAAAAATAATCCAACTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTT
AAGCCTGTCCAAAATTTCGCAGCTGGCTTGTTTAGATACAGTAGTCCCCATCACGAAATTCATGGA
AACAGTTATAATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTT
TTTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATG
CTTTTATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGA
AGAACTTATCCGATTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATT
CATTTGGATTATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAA
CTGTCCTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGT
AGAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAATCG
GGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCTTGCCACT
TTCACCAGCAAAGTTC

SEQ ID NO: 257, cassete de expressão
AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAATAT
AAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCACCTACT
TTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGTAATTAAGT
GGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGTTAAATTATTCG

FIGURA 24 CONT.

AGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTTCTAGCTGAACTCAATGGGT
AAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTCTGAACGTATTGGCAAAGATTTAAA
CATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACATCATTAAGGACATGTCTTA
CTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATTTTTATTATTTATCTTTTTTCG
ATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAAT
ACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTAGGTAGCAATATC
TGAATTCAAGCACTCCACCATCACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTT
ACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAAAAAAGAATT
TTGCTCGTGCGCGAGCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCCACA
GAACAACCCACAAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAG
GCTTTGCGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGCATCCTCCTTCTCCCATCTATAA
ATTCCTCCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAAG
GACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTCGATCCATATCTTCCGGTCGAGTTCTTGGT
CGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCTCCCTTCGGTTGTTCTTGGATTT
ATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCC
TGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTCGGTTTGATTAGTAGT
ATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTAGGGATCGGAATCTTGCGATT
TTGTGAGTACCTTTTGTTTGAGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAGTAC
GGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTC
CCTATTGAACAAAAATAATCCAACTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTT
AAGCCTGTCCAAAATTTCGCAGCTGGCTTGTTTAGATACAGTAGTCCCCATCACGAAATTCATGGA
AACAGTTATAATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTT
TTTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATG
CTTTTATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGA
AGAACTTATCCGATTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATT
CATTTGGATTATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAA
CTGTCCTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGT
AGAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAATCG
GGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCTTGCCACT
TTCACCAGCAAAGTTCATTTAAATCAACTAGGGATATCACAAGTTTGTACAAAAAAGCAGGCTTAA
ACAATGGCGGGGCAGCCGCCGCAGGGCCCGGAGGACGACTTCCTCGACCAGTTCTTCTCCCTCACC
AACTCCCTCTCCGCCGCCGGCCGCCCCTCCGGCGACCAGCCCTTCTCCCTCGCCCTCAGCCTCGAC
GCCGCCTCGGACGCCTCCGGCAGCAGGGGCGGCATCGGCGACGACGCCGGCAACGCCGCGGAGCGG
GACGGCGTGCAGCTCCCCGGCCTCTTCCCGCCGGTGTTCGGCGGTGGCCTCCAGCCGCCGCACCTT
CGCCCCAGCCACCCTCCCCCACAGATGTTCCACGCGCAGCAGCCGAAGCAGGGCGGGCCGGCTGGA
GGGCCGCAGCCGCCGGCGCCGAGGCCGAAGGTGCGGGCGCGTCGCGGTCAGGCGACTGATCCCCAC
AGCATCGCGGAGAGGCTAAGAAGAGAGAGGATAGCGGAAAGGATGAGGGCCCTACAGGAATTGGTC
CCCAATACGAACAAGACAGATAGGGCAGTTATGCTAGATGAGATCCTGGATTATGTGAAGTTCCTT
AGGCTTCAAGTAAAGGTGTTAAGCATGAGCAGATTGGGTGGTGCTGGTGCTGTTGCACAGCTGGTT
TCTGACATTCCACTTTCAGTTAAGGGGGAAGCAAGCGATGGTGGGAGCAAACAGCAGATATGGGAA
AAGTGGTCAACGGATGGCACGGAAAAACAGGTTGCGAAGCTGATGGACGAGGACATCGGTGCCGCA
ATGCAATTTCTCCAATCAAAGGCTCTCTGCATGATGCCAGTCTCCCTTGCCATGGCTATCTATGAC
ACACAACATTCACAGGACGGCCAACCAGTGAAGCCTGAACCCAACAACACTGCCTAG

SEQ ID NO: 258, A. cepa_TC3698

MASNNNPQPPSDGLNDDFFDQIFSMPSAFAASSASTDATEALNYADSSTTGVTPMFSLGLSLDQQP
KSDSSQSEREHPVQFACLFPQYGHNNQVHQTRPNLSPPQVLHGQAMQSSMAATPQPTGPRPRVRAR
RGQATDPHSIAERLRRERIAERIRALQELVPSTNKTDRAVMLDEIVEYVKFLRLQVKVLSMSRLGG
AGAVAQLVADIPMSSSVEGDSSESGKTSYQGWEKWSTDGTERQVAKLMEEDVGAAMQFLQSKALCI
MPISLAAAICQTNPSTEISFM

FIGURA 24 CONT.

SEQ ID NO: 259, A. thaliana_bHLH007-11-AT1G03040
MANNNNIPHDSISDPSPTDDFFEQILGLSNFSGSSGSGLSGIGGVGPPPMMLQLGSGNEGNHNHMG
AIGGGGPVGFHNQMFPLGLSLDQGKGHGFLKPDETGKRFQDDVLDNRCSSMKPIFHGQPMSQPAPP
MPHQQSTIRPRVRARRGQATDPHSIAERLRRERIAERIRSLQELVPTVNKTDRAAMIDEIVDYVKF
LRLQVKVLSMSRLGGAGAVAPLVTEMPLSSSVEDETQAVWEKWSNDGTERQVAKLMEENVGAAMQL
LQSKALCIMPISLAMAIYHSQPPDTSSSIVKPEMNPPP

SEQ ID NO: 260, A. thaliana_bHLH059-11-AT4G02590
MASNNPHDNLSDQTPSDDFFEQILGLPNFSASSAAGLSGVDGGLGGGAPPMMLQLGSGEEGSHMGG
LGGSGPTGFHNQMFPLGLSLDQGKGPGFLRPEGGHGSGKRFSDDVVDNRCSSMKPVFHGQPMQQPP
PSAPHQPTSIRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPTVNKTDRAAMIDEIVDYV
KFLRLQVKVLSMSRLGGAGAVAPLVTDMPLSSSVEDETGEGGRTPQPAWEKWSNDGTERQVAKLME
ENVGAAMQLLQSKALCMMPISLAMAIYHSQPPDTSSVVKPENNPPQ

SEQ ID NO: 261, A. thaliana_bHLH066-11-AT2G24260
MMNSSLLTPSSSSSSHIQTPSTTFDHEDFLDQIFSSAPWPSVVDDAHPLPSDGFHGHDVDSRNQPI
MMMPLNDGSSVHALYNGFSVAGSLPNFQIPQGSGGGLMNQQGQTQTQTQPQASASTATGGTVAAPP
QSRTKIRARRGQATDPHSIAERLRRERIAERMKALQELVPNGNKTDKASMLDEIIDYVKFLQLQVK
VLSMSRLGGAASVSSQISEAGGSHGNASSAMVGGSQTAGNSNDSVTMTEHQVAKLMEEDMGSAMQY
LQGKGLCLMPISLATAISTATCHSRNPLIPGAVADVGGPSPPNLSGMTIQSTSTKMGSGNGKLNGN
GVTERSSSIAVKEAVSVSKAATGMQFDESDARVWIWFEIRREDDIVELLWQSGQVVGTNQTHRQSY
DPPPILRGSGSGRGEENAPLSQPPPHLHQQNLFIQEGEMYSWLHHSYRQNYFCSELLNSTPATHPQ
SSISLAPRQTIATRRAENFMNFSWLRGNIFTGGRVDEAGPSFSVVRESMQVGSNTTPPSSSATESC
VIPATEGTASRVSGTLAAHDLGRKGKAVAVEAAGTPSSGVCKAETEPVQIQPATESKLKAREETHG
TEEARGSTSRKRSRTAEMHNLAERRRREKINEKMKTLQQLIPRCNKVESDSVSTLISLLKFQRWMM
LSSTSNRYRAKYKYALQNRMCFKPMVQHGKSSYVSSFVEMMSTGQGMMSPMMNAGNTQQFMPHMAM
DMNRPPPFIPFPGTSFPMPAQMAGVGPSYPAPRYPFPNIQTFDPSRVRLPSPQPNPVSNQPQFPAY
MNPYSQFAGPHQLQQPPPPPFQVTLYHGIHCVVAGNPYV

SEQ ID NO: 262, A. thaliana_bHLH069-11-AT4G30980
MNSSSLLTPSSSPSPHLQSPATFDHDDFLHHIFSSTPWPSSVLDDTPPPTSDCAPVTGFHHHDADS
RNQITMIPLSHNHPNDALFNGFSTGSLPFHLPQGSGGQTQTQSQATASATTGGATAQPQTKPKVRA
RRGQATDPHSIAERLRRERIAERMKSLQELVPNGNKTDKASMLDEIIDYVKFLQLQVKVLSMSRLG
GAASASSQISEDAGGSHENTSSSGEAKMTEHQVAKLMEEDMGSAMQYLQGKGLCLMPISLATTIST
ATCPSRSPFVKDTGVPLSPNLSTTIVANGNGSSLVTVKDAPSVSKP

SEQ ID NO: 263, A. thaliana_bHLH082-11-AT5G58010
MENGNGEGKGEFINQNNDFFLDSMSMLSSLPPCWDPSLPPPPPPPQSLFHALAVDAPFPDQFHHPQ
ESGGPTMGSQEGLQPQGTVSTTSAPVVRQKPRVRARRGQATDPHSIAERLRRERIAERMKSLQELV
PNTNKTDKASMLDEIIEYVRFLQLQVKVLSMSRLGGAGSVGPRLNGLSAEAGGRLNALTAPCNGLN
GNGNATGSSNESLRSTEQRVAKLMEEDMGSAMQYLQGKGLCLMPISLATAISSSTTHSRGSLFNPI
SSAVAAEDSNVTATAVAAPEASSTMDDVSASKA

SEQ ID NO: 264, A. vulgaris_TC15501
MANNNNPTEGLTDDFLDQILSYGAHEGNLAGNDVNLAGTTTPGGSMALQLSSSGGDITGSGGYHHH
HQHQHHHQLSGGGGGFPLGLSLEQSGKGGFFKPDEASGSGKRFRDDFVDLKAVSSTNNDNNDRGDS
VHLNNLFPAFGHVQQQQQAHSVRPPSHQNQINQPFLGHPTPGAVAVVPHPPAIRPRVRARRGQATD
PHSIAERLRRERIAERMKALQELVPSSNKTDRAAMLDEIVDYVKFLRLQVKVLSMSRLGGAGAVAQ
LVADIPLSSAEGGASEGGSNQPAWEKWSTDGTEHQVAKLMEEDVGAAMQFLQSKALCIMPISLASA
IYHDQPPDASTMIKPEPNAPS

FIGURA 24 CONT.

SEQ ID NO: 265, A. vulgaris_TC19278
MKNLQELVPNSNKTDKASMLDEIIEYVRFLQLQVKVLSMSRLGATGAVVPLITENQPEGSTSHFLS
QSAGRDVKHIESPNTLTFEQEVVKLMESNMTKAMQFLQRKGLCLMPIALAAAISGSKAPPASLSTE
GKKSILTNGFVHHNSGQSNIGPGQISSDGETRSEENIIGIHSREDFSSNSCSGTSIKKEGTNTTNF
TREMNQE

SEQ ID NO: 266, B. napus_TC10015_part
HASETPSDDFFEQILGLPNFSASSSDGGLGGGGAPPMMLQLGSGEEGGHMGGLGGGSGFHNQMFPL
GLSLEQGKGQGFLRPEGGSLGTGKRFSDDVMKPVFHGQPMQQQPAPAAPHQPTSIRPRVRARRGQA
TDPHSIAERLRRERIAERIRALQELVPTVNKTDRAAMIDEIVDYVKFLRLQVKVLSMSRLGGAGAG
APLVTDMPLSSSVEDESGEGGRAPQPAWEKWSNDGTERQVAKLMEENVGAAMQLLSSKALCMMPIS
LAMAIYHSQPPDTT

SEQ ID NO: 267, C. clementina_DY268946
MFLQLGSGGPSPGAGGGATSAADIRSLAMGISMGMMPLGLNLEHSFLRQHEDHNSNHNNNNNTNAS
SSSPNNSSATSGINERDSVHMPSLFPAFGQLQSQSARPPLPRPPQVQQFQSQPAAAPQPPAVRPRV
RARRGQATDPHSIAERLRRERIADRMRALQELVPSCNKTDRAAMLDEIVDYVKFLRLQVKVLSMSR
LGAAGAVAQLVADVPLSSALEGESIDGGSSQPEWEKWSNDGTEQQVAKLMEEDIGAAMQFLQSKAL
CIMPISLASAIYRTRQPDAPAFVKPESSTPSYCLNNNWSQQLAHNDSKYLITGTYSYRILLVLLSN
QIIQIPMQRSHSSFP

SEQ ID NO: 268, C. longa_TA2544
MQPSSKETERMAQSHAALQLELQNGGGGAQNDDFFEQTMSGFSTAAWAPELGTPRSLFGARSSEES
PDDEGLNYAPPYGGSSLLASRHRMHLGTSPADSSMFLQLGGQILLHPTAGASGGESGGFLRLPLSL
GSGGSGVSTSAIFGDRSRGEIDPPFESSNPTEAEVLYNDGFTGSLPRAVQASLHQQLLRQPQNYGA
ATLTGQALAITATASGTAGAGTAPPKPRVRARRGQATDPHSIAERLRRERIAERMKALQELVPNAN
KTDKASMLDEIIDYLKFLQLQVKVLSMSRLGGAAAVAPLVADMSSESGGGAGTPRSNNDGLTAAEQ
QVAKLLEEDMGSAMQYLQGKGLCLMPISLASAISSATCRPRPPPGVNLRYPAAGDAPPSPSISALT
VQSSNAGSAGADAS

SEQ ID NO: 269, C. paradisi hybrid_TA3392
MANNPNEASSTDDFLEQILGIPNFGSAESGLAASDGGLAAGSPMMLQLSSGDGSSHISALGGGVSS
GYHGQVFPLGLSLEQGKGGFLKPEEASGSGKRFPEEHAIKNVFHGQPLPSPVPAAPHPPAMRPRVR
ARRGQATDPHSIAERLRRERIAERIRALQELVPSVNKTDRAAMLDEIVDYVKFLRLQVKVLSMSRV
GAPGAVAPLVTTDLPLSSVEDESGEGVRNQPAWEKWSNDGTERQVAKLMEENVGAAMQFLQSKALC
IMPISLATAIYHSQPPE

SEQ ID NO: 270, C. sinesis TA12416
MQPCSSGGEMQGISSLLSNGSHEQQSQIHQSATFDPTSQDDFLEQMLSSLPSCSWTDLKSPWGVVD
LNPNNNNNNINNKQPRDLSDETAPSTTQENNVPAGFQFDESMILASKMRQHQISGNTGSGNNNSAA
AKFMMQQQQQQIMMAAARGGIGLPLSLGNGGDNDIVDVSSFKSQQGGDGSVQALYNGFTGSLHGST
QPQHFHHLQGGSMPGQTFGAPGPVMNQTQAQASGSTGGGGGGGNTPAQQPKQRVRARRGQATDPHS
IAERLRRERIAERMKALQELVPNANKTDKASMLDEIIDYVKFLQLQVKVLSMSRLGGAAAVAPLVA
DMSSEGGGGDCIQANGRNPNGAQTTSANDSLTVTEHQVAKLMEEDMGSAMQYLQGKGLCLMPISLA
TAISTATCHSRNPIISTSNNNNNNGNPHHNPLLQSNGEGPTSPSMSVLTVQSATMGNGGADGSVKD
AASVSKP

FIGURA 24 CONT.

SEQ ID NO: 271, E. grandis WO2005001050 seqID77
MAGSNPQEGLGDEFFEQILAVPPGAYSGGSTPMVLQLGSGRGGVEDGGGRGGGGGLRGMGVMMPLG
LNLEQGGLFRHEDVENSTSASSTTSAINMEREAGVHHGNNMTSCLFPAFGQFQTHQSVQPPPPHPP
PPQLHPAFLNQPAVGSPNQPAIRPRARARRGQATDPHSIAERLRRERINERMKALQELVPSCNKTD
RAAMLDEIVDYVKFLRLQVKVLSMSRLGGAGAVAQLVADVPLSSAEGEIIEGGNNQPAWEKWLTDG
TEQQVAKLMEEDVGAAMQFLQSKTLCIMPISLASAIFRTSQPDMPRSIKPESSAPS

SEQ ID NO: 272, E. grandis WO2005001050 seqID1671
MANNPSDGASDEFLEHLLGLPNFAAAAEAGLAPGDGTGLSVTAATPMMLQLGSGDGSGGHGHGHGH
GHGQGHLSALGGGGAGAGGGGGGGGGGYHGAVFPLGLSLEQGKGGFLKPEEASGSGKRYPEEVVDG
RASTVKNLHSPLPNFPDVSQMASFRNRDVFHGQPVPNPVPAAPHPPAMRPRVRARRGQATDPHSIA
ERLRRERIAERIRQLQDLVPSVNKTDRAAMLDEIVDYVKFLRLQVKVLSMSRLGAAGAVAPLVTDI
PLSSVEEEGGEGGRNQPAWEKWSNDGTERQVAKLMEENVGAAMQFLQSKALCIMPISLATAIYQTQ
PPDTSSLVKSESNPPS

SEQ ID NO: 273, G. hirsutum CO123623
DSGGGGFRGIGMMPLGLNLEHGFLRHEDGVVVDNNNNNASCSAASAVSGISERDSMHMVSLFPPFG
QMQTQQIRASPPPPQPPPQLHQPFHSQPTSGPVAAAPHPPAVRPRVRARRGQATDPHSIAERLRRE
RIAERMKALQELVPSCNKTDRAAMLDEIVDYVKFLRLQVKVLSMSRLGAAGAVAQLVADVPLLSVE
GDGAGGGTQPPWEKWSNDGTEQQVAKLMEEDIGAAMQFLQSQALCIVPISLASAIFPAHQPDAPTI
V

SEQ ID NO: 274, G. hirsutum DT543504
MISNGGESSEFRRSFTGLETLSPIPQLWHLQPYDSVSSLPTLVGQTIVDDEGNINRFIEIDEILQP
ENLSASINSKGQQDMQNSFCSSFPADHPITKTMIGLPSLVQGASPNLNNGSDRTVKPRVRARRGQA
TDPHSIAERLRREKIAERMKNLQELVPNSNKTDKASILDEIIGYVKFLQLQVKVLSMSRLGAAAAV
VPLITDGRAEVSNGLSLAPLAGQGVDFSPSPDQVVFEQEVVKLMESNMTMAMQYLQSKGFCLMPVA
LAAAISNVKSSTSSSSSSVPASDESKKLGLHQQYSNQQHLQQ

SEQ ID NO: 275, G. hirsutum TC60118
MANNPNESPADDFLEQILGLSHFAPTETGLAGPDGRLSGNATTAGAPMLLQLSSGGGTGHIGAIGG
GGGGAFHGQVFPLGLSLEQGKGGFLKPEEASGSSKRFRNEVVDGRAFSVKNVFHGQPVPATVAAGP
HPPAMHPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVNKTDRAAMLDEIVDYVKFLRL
QVKVLSMSRLGAAGAVAPLVTDLPLSSVEDESGEGGRSQPAWEKWSNDGTERQVAKLMEEDVGAAM
QFLQSKALCVMPISLATAIYHTQSPDTSSVVKPETNPPS

SEQ ID NO: 276, G. hirsutum TC60119
MANNTNEAPAADDFLEQILGLPNFAPSETGLAGSDAGLAATAAGAGAPMFLQLSSGDGAAHIGGIG
GGGGGAFHGQVFPLGLSLEQGKGGFLKPEEASGSGKRFRNGVVDDRASSVKNVFHGQPMQATVSAA
PHPPTMRPRVRARRGQATDPHSIAERLRRERITERIRALQELVPSVNKTDRAVMLDEIVDYVKFLR
LQVKVLSMSRLGGAGAVAPLVTDIPLSSVEYESGEGGRSQPAWEKWSNDGTEQQVAKLMEENVGAA
MQFLQSKSLCIMPISLATAIYHTQVPDTSSVVKPETNPPA

SEQ ID NO: 277, G. hirsutum TC61833
MANNPNEASADDFLEQILGFPNFAPSETGLAGPDGGLTGTTAGAGTPMFLQLSSGDGASHLGGIGG
GGGGAFPGQVFPLGLSLDQGKSGGFLKPEEGSGGSSRRFRDEVVDGRASSVKNVFHGSPMPATVAP
SPHPPTMRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVNKTDRAVMLDEIVDYVKFL
RLQVKVLSMSRLGGAGAVAPLVTDIPLSSVEDESGDGGRNQPAWEKWSNDGTERQVAKLMEENVGA
AMQFLKSKALCIMPISLATAIYHSQPLDTSSIVKPETDPPP

FIGURA 24 CONT.

SEQ ID NO: 278, G. hirsutum TC67603
MQPCSREMQAMNSLLNPTQQIPLQDLQINGNRHHHQQIHVPSSSQFQYPTPTSTTHHDDSFLEQIL
SSTTSFPWSDETGPPNPEDNNNAGFNYDEMVLASKLRQHQINGGAAGNPFAAMKMMLMMQQQQQQQ
QMMLAGRPTVASAAAGGGITTNHQGGEGSMQALYNVFGDGSLHGTMQAKSFVGANSATEMTQSQGS
GPTAGEAVTAPEKPKQRVRARRGQATDPHSIAERLRRVRIAERMKALQELVPNANKTDKASMLDEI
IDYVKFLQLQVKVLSMSRLGGAAAVAPLVSEGGGDCIQTSANGGSHPRNSNGDQTPSANDRLTMTE
HQVAKLMEEDMGSAMQYLQEKGLCLMPISLATAISSATSRSRNPMINHENPANGSHLLIQSSGGDG
PSSPSMSVLTVQSATMGNGGLEGGAASVSKP

SEQ ID NO: 279, G. max TC229602
MKPDEASASGKRFRDDVVDNRAKHVFHGQPMPTTMPAAPHPPAIRPRVRAIRGQATDPHSIAERLR
RERIAERIRALQELVPSVNKTDRAAMLDEIVDYVKFLRLQVKVLSMSRLGGAGAVAPLVTDIPLSS
VEEEGGEGARNRPAWDKWSNDGTERQVAKLMEENVGAAMQFLQSKALCIMPISLASAIYQSQPPDT
SSIVKPETNPPS

SEQ ID NO: 280, G. max TC205173
MLQHQLLRDSGLLNMPLSLPGNDVVVDASTFESPNPSGKASVQTLYNGFAGSLHGVGQSSNQTQHF
QHPQGSSNPMQGQNFGAVPAGGGSATNQAPASGAAAGGAPAQPRQRVRARRGQATDPHSIAERLRR
ERIAERMKALQELVPNANKTDKASMLDEIIDYVKFLQLQVKVLSMSRLGGAAAVAPLVADMSSEGG
GDCIQANGKSNGGGAQASTTNTNTNQTTATTTSNDSLTMTEHQVAKLMEEDMGSAMQYLQGKGLCL
MPISLATAISTATCPTRNVNVNPLINAAAGATHFPTAANPAGEGPSSPSMSVLTVQSAIAGNDGAA
SVSKP

SEQ ID NO: 281, G. max TA55042
MAGNSGSEGLGDDFFEQILAVPEAGTVGMLQLGSTTGAFRGASGLMPLGLNLEQAAFLRHQVNVDD
DVVHVNVDDATIHQHHLTLHNNNNSSSPSSTAPITDRDSMHMRGLFSAFGQLHTPIRPTLPLPPPP
RQPQLHLHHHNQFQGQAAAAPASMAAMPQPPGIRPRVRARRGQATDPHSIAERLRRERIAERMKAL
QELVPSINKTDRAAMLDEIVDYVKFLRLQVKVLSMSRLGGAGAVAQLVADVPLSAVEGDQDIEGGA
NEQAWDKWSNDGTEQQVAKLMEEDVGAAMQFLQSKALCIMPISLASAIFRMPQSEASTGIKPESNS
H

SEQ ID NO: 282, G. max TC207545
MMLQLNSGDAATHLASFHAPPYQLGLSLDQGEGPFLTPEDASGSGKRFRDDAVDTRPNNVFDGQPM
PTTVPTAPHPPAVRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVNKTDRAAMLDEIV
DYVKFLRLQVKVLSMSRLGGAGAVAPLVTDIPLSSVEEEGGEGRNQPAWEKWSNDGTERQVAKLME
ENVGAAMQFLQSKALCIMPVSLASAIYQSQPSGTSSIVKPETNPPS

SEQ ID NO: 283, G. raimondii TA13791
MANNPNESPADDFLEQILGLSHFAPTETGLAGPDGRLSGNATTAGAPMLLQLSSGGGTGHIGAIGG
GGGGAFHGQVFPLGLSLEQGKGGFLKPEEASGSSKRFRNEVVDGRAFSVKNVFHGQPVPATVAAGP
HPPAMRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVNKTDRAAMLDEIVDYVKFLRL
QVKVLSMSRLGAAGAVAPLVTDLPLSSVEDESGEGGRSQPAWEKWSNDGTERQVAKLMEEDVGAAM
QFLQSKALCVMPISLATAIYHTQSPDTSSVVKPETNPPS

SEQ ID NO: 284, H. petiolaris TA1140
MSTDPPEGYGDDFLEQILAIPSYNITGCLPVNTADSTGSETVSVHHQQQQPQQQLQPQKFPLGLSL
DNGRETIGGAFAGQQQQLQQQQQRERGGSMNMSGLFPQFENLQSHSLLHTVPQGFQGQPTSSTAVT
VPHPPTIRPRVRARRGQATDPHSIAERLRRERIAERMRALQELVPSCNKTDKAAMLDEILDYVKFL
RLQVKVLSMSRLGGAGAVAQLVSDVPLQSVEGDGSENGYNNQQQGQAAWENWSNDDTEREVAKLME
EDVGAAMQFLQSKALCIMPISLASLIYPNHQPDVKPEPSAPS

FIGURA 24 CONT.

SEQ ID NO: 285, H. vulgare TC140470
MNYSDGSFFPSWPGNSASENYSFVDGSVESYTEEGSMPTTGYFRARSNQNLTFDEHEQNPAMLANG
CLPYNTQTDLLSGEILSDDKPSNSLVELSQLQNNVCLQNNLIPPGTLQCNSTPGTFDLQLDTPGLL
ELPHALSSSIESNGSEVSAFLADVQAVSSASTLCSTFQNVPSYMEPVSLEAFSFQGMQNAAMFNNA
SHSNGNLSVFDEATMASLHDSKEFLNGSISSLGTGQQSQLAGGGLKAEQQEQNTMCNIPLPSFVSA
SQMAVSEAQGALIPSKTTSITHNNKSEYPIPISHSADVQHKANSGNGNSASAKPRARARRGQATDP
HSIAERLRREKISERMKNLQDLVPNSNKADKSSMLDEIIDYVKFLQLQVKVLSMSRLGAPGAVLPL
LAESQTEGRSNSPLSSPTTSQGLLDVAGPEDSLVFEQEVIKLMETSITNAMQYLQNKGLCLMPIAL
ASAISNQKGTSAAAIPPER

SEQ ID NO: 286, L. perenis TA3490 part
MATDPPEGYGDDFLEQILAIPSYNIAGCLPGNTTDANASETVSVHRQQQQQQPVFPLGLSLDNGRE
TIGAFAGQQQQRERGGSMNMTGLFPSFENLQSHSLLHSVPQAFQGQSTTSTAVTVPHPPSIRPRVR
ARRGQATDPHSIAERLRRERIAERMRALQELVPSCNKTDKAAMLDEILDYVKFLRLQVKVLSMSRL
GGAGAVAQLVSDVPLQSVEGDTSENGYNQPAWENWSNDDTEREVAKLMEEDVGAAMQFLQSKALCI
MPISLASLIYPPQTPDSNSHVKPETV

SEQ ID NO: 287, N. benthamiana CK293938
MATNQPEGYADDFLEQILAIPSYAGLPPVADVGTSSETTSFTSASAAGLQQPLFPLGLSLENGRDD
VSDAGAYAVKHEREGINIGNLYAGLEHLQSHAVRHAVPPVHHIQSFQGQPTTSTTMTVPHSPAIRP
RVRARRGQATDPHSIAERLRRERISERIKALQELVPSCNKTDRAAMLDEILDYVKFLRLQVKVLSM
SRLGGASAVAQLVADIPLQSMEGDSAESKSNQHIWEKWSNVDTEQEVAKLMEEDVGAAMQYLQSKS
LCIMPISLAALIYPTQQPDDMSMVKPEPAAPS

SEQ ID NO: 288, N. benthamiana TC7102
MQAMNQYDPTSSHDDFLDQILSSVPSSSLCWPDLSKSWDPHHHHLSPPLPPNPSSGDDHQLPPSNP
LQHFQYDDQSSSFLAAKLRQHQITGGGGGGGDAVAAAKALLLQQQLLLSRTLAGNGLRSPTGASGD
NGFLNMLGNGDQNDSVGNPANDSSVQALFNGFTGSLGQNSSQPQHFLHPQGGRMQAQSFGAPATAP
AMNQTPAASGSAGGGTTPAAQPKQQRVRARRGQATDPHSIAERLRRERIAERMKALQELVPNANKT
DKASMLDEIIDYVKFLQLQVKVLSMSRLGGAAAVAPLVADMSSEGRGEGNGGRGGNGTASSSNNDS
MTVTEHQVAKLMEEDMGSAMQYLQGKGLCLMPISLATAISTATCHSRNPMIPNGNGGNNPLLGGGG
GLATNGGGGEAGGPSSPKLSVLTVQSATIGNGGIDPSVKDATSVFEA

SEQ ID NO: 289, N. benthamiana TC7103
MQAMNSQMSLQEENGASSGQNMSHSHFDPTSSHDDFLQQILSSVPSSSPWPDLSSGGDGHFDDQSI
FVASKLRQNQITGGGGGGAAAKALMLQQQLLLSRGLLAGNGDQNDDVGNPENEISVQALYNGFAGS
LGQTSNQPQHFHHAQGGSMQAQSFGAPASSTMNQTPAASGGSAGGGQPKQQKVRARRGQATDPHSI
AERLRRERIAERMKSLQELVPNANKTDKASMLDEIIDYVRFLQLQVKVLSMSRLGGAAAVAPLVAD
RSSEGGGGDCAQANGGGRGSNGTTSLANNDSMTMTEHQVAKLMEEDMGSAMQYLQGKGLCLMPISL
ATAISTATCHSMKLNNPLLHGGGSGGSAAINGGGGGPSSPSLSASTVQSATMGNGGA

SEQ ID NO: 290, N. benthamiana TC8633
MANNPSEGPDDFFDQILGFPAYNGAETNLAGSDAGAIPPAMLLQLNSGDASGQFSGMGLGVGLGGG
GGFHHPSQSGSFPLGLSLEGGKSGFMKMDDVPAVPGRRFRDDVVDSRASSSVKPGFHGQPMPSMPH
PPAIRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPNVNKTDRAIMLDEIVDYVKFLRLQ
VKVLSMSRLGGAAAVAPLVTEIPISSVEEEISEGGNNQPAWEKWSSDGTERQVAKLMEENVGSAMQ
FLQSKALCIMPISLASAIYHSQPPDTSSLIKPETNPPS

FIGURA 24 CONT.

SEQ ID NO: 291, N. tabacum TC12771
MATNQPGGYADDFLEQILAIPSYAGLPPVADVGTSSETTSFTSSSVAAAGLQQPLFPLGLSLDNGR
DDVSDAGAYAVKHEREGINIGNLYAGLEHLQSHAVRHAVPPVHHIQPFQGQPTTSTTVTVPHPPAI
RPRVRARRGQATDPHSIAERLRRERISERIKALQELVPSCNKTDRAAMLDEILDYVKFLRLQVKVL
SMSRLGGASAVAQLVADIPLQSVEGDSAESRSNQHIWEKWSNVDTEQEVAKLMEEDVGAAMQYLQS
KSLCIMPISLAALIYPTQQPDDPSMVKPEPAAPS

SEQ ID NO: 292, O. sativa Os02g35660
MGGFAYPFTPSPAWSRDAVFAGSPWAAGGVSSLADALVSYGAVDDEEAAFLGKTAASSPSTARLHE
QQQLLLEAELLRHGDGLGFAAMDDDGGAAMLGALEPCAMPLTDSGGPPVICSSSSNDSSGSEHSAA
MPAGGGFLVGEQQQHVPPAAYAAGGVLPSMAAGEETPQSFGFGSLFNGDLLQEATVSKYHHHQQQQ
QLGVVPSSQPHHLNDDIDFNTGKLMSFASGQQHVTPSIDSLQIDQKEFSSGLHHLNLSSLISGPLA
SFNATQSHRQPAEACGGKNGGAAPFVNLSEVLPKGNGSGSAGNGAPKPRVRARRGQATDPHSIAER
LRREKISDRMKDLQELVPNSNKTNKASMLDEIIDYVKFLQLQVKVLSMSRLGAAEAVVPLLTETQT
ESPGFLLSPRSSSGERQAGAGAVTGGLPGDQPELLDGGAMFEQEVVKLMEDNMTTAMQYLQSKGLC
LMPVALASAISAQKGTSSAAVRPEKKKNGDGDGGGDEEDVKGEFDAPRRPPVGRPKEMRSRV

SEQ ID NO: 293, O. sativa Os02g55250
MQPNARQMRGGGGGGGGQDDFFDQMLSTLPAVWSELGSGKPAWDLTAGAVGGGGGASDDHSAAAF
DDSALLASRLRQHQIDGGGDKPIMLQLSDLHRHHGLAAGDDSGGAAGFLPLSLFADRSQDDIDAAF
KSPNGARGDHALYNGFGAAGMHGAAAMQPPPFGQGGSMPAQSFGGGAAASGGGGGGSASAAAAAGA
SSGGGAAAPPRQRQRARRGQATDPHSIAERLRRERIAERMKALQELVPNANKTDKASMLDEIIDYV
KFLQLQVKVLSMSRLGGASAVAPLVANMSSESNGNGNATSSSGNGEAANGSSNGDNNGGGTLRVTE
QQVAKLMEEDMGSAMQYLQGKGLCLMPISLATAISSATSSSLLPRTGGGAGGSLHEGGNGTSPPLV
NGTATGCDDAGVFFSVKFVVELSFLLLNEDCRGKEESKLLVQKGP

SEQ ID NO: 294, O. sativa Os03g58330
MAGQQPQQQGPPEDDFFDQFFSLTSSFPGAAPGGRAAGDQPFSLALSLDAAAAAEASGSGKRLGVG
DDAEGGGSKADRETVQLTGLFPPVFGGGGVQPPNLRPTPPTQVFHPQQSKQGGAAVGPQPPAPRPK
VRARRGQATDPHSIAERLRRERIAERMRALQELVPNTNKTDRAAMLDEILDYVKFLRLQVKVLSMS
RLGGAGAVAQLVADIPLSVKGEASDSGGNQQIWEKWSTDGTERQVAKLMEEDIGAAMQFLQSKALC
MMPISLAMAIYDTQQTQDGQPVKHEPNTPS

SEQ ID NO: 295, O. sativa Os06g08500
MQPSSRDTVAGGGGEGTQDDFFDQMLSTLPSAWADLGGGGGGAAGKSPWEVDPAAAAAASQVFDES
ALLASRLRHHQIGGAGGGGGEKPVMLQLSELHRQAGGGEEDGSGAFSPLPLFTDRTNVPPREEMEG
GFKSPNAAAGGEHALFNGFGVHGGSGGAGQPPFGQLRRERIAERMKSLQELVPNANKTDKASMLDE
IIDYVKFLQLQVKVLSMSRLGGAAGMAPLVASMSSEGNSNGSSNGGGGKASKGGTGGEGGGGGGGG
GGGGTGGGMRVTEQQVAKMMEEDMGTAMQYLQGKGLCLMPISLASAISSATSSASLLSRPSIRHAG
APPQTMLDAAGPTSPAAMSNGDDPRHAKADGGAGGTQ

SEQ ID NO: 296, O. sativa Os06g09370
MDYSAGSYMWPGNSGSENYNFVDGSSESYAEEGSLPPSGYFMGAGSDRSLKITENERNPTMLANGC
LPYNTQAHPLSGQILPKGELPNNLLDLQQLQNSSNLRSNSIPPGVLQCNSTSGTFDAKLDTPGLAE
LPHALSSSIDSNGSDISAFLADVHAVSSAPTLCSAFQNVSSFMEPVNLDAFGFQGAQNVAMLNKTS
LPNGNPSLFDNAAIASLHDSKEFLNGGSIPSFGTVLQALGAGGLKAAQQEQNIRNIPLPTFTSGSH
LAVTDAQGPPLPSKIPPLIHDHNSEYPINHSSDVEPQANSAPGNSANAKPRTRARRGQATDPHSIA
ERLRREKISERMKNLQVLVPNSNKADKASMLDEIIDYVKFLQLQVKVLSMSRLGAPGAVLPLLRES
QTECHSNPSLSASTISQGPPDMPDSEDSSAFEQEVVKLMETSIISAMQYLQNKGLCLMPIALASAI
SNQKGMAAAAAIPPEK

FIGURA 24 CONT.

SEQ ID NO: 297, O. sativa Os09g25040
MAAAAAAGQISLDDLRNGGGVAANAGGGGGVHDDFLDQMLSSLPPSAWPDLAAGKAAEDDAEGMHH
HHHQQQQQFGGPYDESAMLASRLRQHQISGGGGGGGGGAAAVKQMVLQQLADLRQGHHMMLQGLGG
RSPAGGGGGGGDGGLLLPLTLGSGGSGGDVQALLKAAAANSAGGGDAGGVYGGGFAGSLHQQQQHF
QPHPQTAPTIPTQSFGGGGGGGGGGTASGGGAAQPQAGAAGGGAPAPPRQRVRARRGQATDPHSIA
ERLRRERIAERMKALQELVPNANKLMQTDKASMLDEIIDYVKFLQLQVKASTYTKLLIHVLSMSRL
GGAAAVAPLVADMSSEGRGGGAANGGAPAAAGSDSLTVTEQQVAKLMEEDMGTAMQYLQGKGLCLM
PISLASAISSATCHLRPPVVAAAQQFPAGLGAAAAAAHHHQLSAAAAAAAMRGHLPGLNADGSVPA
SPSMSVLTAQSAMANGGGGAADGEGSQLKDAASVSKP

SEQ ID NO: 298, O. sativa Os07g08440
MAGQQPPTGGAEDDFFDHFFSIPSAAAAGAGGVGGFGSGDHHPFPLALSLDAEGAGAARRLLDGGH
DGGRTDRDPVQLAGLFAPVFGAAAGVQPPHLRAPPPPQVFHAQPKPGEGAMAAPQPQQPPAPRPKV
RARRGQATDPHSIAERLRRERIAERMRALQDLVPNTNKTDRAAMLDEILDYVKFLRLQVKVLSMSR
LGGAGAVAQLVADIPISVKGEASDSGSKQQIWEKWSTDGTEKQVAKLMEEDIGAAMQFLQSKALCM
MPISLAMAIYDTQHSQDGHSVKPEPNTPS

SEQ ID NO: 299, P. abies DV982110
MLPLPLSLGQAGKSGDMNEPGSREEIEASFKSVNNARDSNLGGLFQPFAVSPRGVRPTGQNFHAQP
GQVPLQGYGGMPQPQHQNPPPTGVGAAPPVRPRVRARRGQATDPHSIAERLRRERIAERMKALQEL
VPNSNKTDKASMLDEIIDYVKFLQLQVKVLSMSRLGGAGAVAPLVADIPAEGANAPQGTRTNGSQN
SSPDGLALTERQVAKLMEEDMGTAMQYLQGKGLCLMPISLASAINNSGSRPQAPSTPSXXXLLASN
VNDRQVLEPS

SEQ ID NO: 300, P. deltoides TA3263
MANNPTEPPADDFLQEILGMPNFASAEAGLVGADAGLAGAAAAQASMMLQLSSGDGSGHISDLGGA
PGGGSAGFHGFPLGLSLEQGKGGFLKPEEASGSGKRFRDEIVDGRAKNVFHGQPMPTTVAIAPHPP
AMRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVNKTDRATMLDEIVDYVKFLRLQVK
VLSMSRLGGAGAVAPLVTDIPLSSVEDETGEGGRNQPAWEKWSNDGTERQVAKLMEENVGAAMQFL
QSKALCIMPISLATAIYHTQPPDTTSIVKPETNPQS

SEQ ID NO: 301, P. radiata WO2005001050 seqID516
MSLLPPGIGGQGLNSQENDPYPYDDSQFLANRLRQHQLSGNSSMNQTPNQASQGINETNTSPGRSM
VLQLSTGSASASQLLASMGNSPRGTAVMTQSPSTGGSCNGSDGGMLPLPLSLGQAGKSGDINESRS
RDEIEASFKSANHARDSNLGGLFPPFAASPRGVRPTGQNFHTQSGQVPMQVYGGMPQPQHQNPSPT
GVGAAPPVRPRVRARRGQATDPHSIAERLRRERIAERMKALQELVPNSNKTDKASMLDEIIDYVKF
LQLQVKVLSMSRLGGAGAVAPLVADIPAEGANAPQGTRTNSSQNSSPDGLALTERQVAKLMEEDMG
SAMQYLQGKGLCLMPISLASAINNSGARPQAPSTPSLQGLLPPNANDRQILEPSNSALSALTSTSM
TIQSSISSPGSGITEQGDNAHKAVNGTRNPKDSREANSVTKSNGIGGPALSKNAVKGEDEPQRRTG
Q

SEQ ID NO: 302, P. sitchensis TA14134
MDDIFDQLLSSSWEDINGADRSSLDSSAGGPTNCLPGDSMGTLQASSRVSTMSVTPSSHMSPNLEG
QAQNNDVQNSSSSVIAGENTPYLLKELLVCSQATDPHSIAERERRERIAKNLKSLQELVPNANKTD
KASMLDEIIDYVKFLQLQVKVLSMSRLGGAGAVAPLIADVPAEGSGSLASAALGQAGGSLSQDGLA
FEQNVVKLMEKDMTAAMQYLQNKGLCLMPIALATAISSSTGKPLLGSGVTNSVDSTSDRQSSGNQP
ASLTVDSCSGALGMGSAGFGSDFLLKENTVEKRGRELVPTAESNGAEFGIISSLPKLRKKE

FIGURA 24 CONT.

SEQ ID NO: 303, P. taeda TC68930
MLARENRSGDTEHDDLQGTLLTSYTGPQGGQTVLPRTPGAQSHQQNLSTQTIQSGEGTSLQQYGNH
SAVSQLQSGAGGGNAANGAARPRVRARRGQATDPHSIAERLRRERIAERMKSLQELVPNSNKTDKA
SMLDEIIEYVKFLQLQVKVLSMSRLGGAGAVAPLIADVPSQGSGGVMSTALGQASGPLTLSQDGLA
FEQEVARLMESNMTSAMQYLQNKGLCLMSIALATAISSSSGKPVLATVPGTGVDSSEKQNSDMQSI
VLPFSSSSTTMGLRATASGSDAPINENSINKASIEKVVAEKSNGISPGLSSDVPKVGSQSREELLK
RAQ

SEQ ID NO: 304, P. trichocarpa scaff II. 416
MANNPTEPPTDDFLQEILGMPNFASAEAGLVGADAGLAGTASVQAPMMLQLSSGDGSGHISALGGA
PGGGGAGFHGFPLGLSLEQGKGGFLKPEEASGSGNRFRDDIVDGRVRNVFHGQPMPTTVTAATHPP
AMRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVNKTDRAAMLDEIVDYVKFLRLQVK
ILSMSRLGGAGAVAPLVTDIPLSPVEDETGEGGRNQLAWEKWSNDGTERQVAKLMEENVGAAMQFL
QSKALCIMPITLATAIYHTQPPDTTTIVKPETNPPS

SEQ ID NO: 305, P. trichocarpa scaff 70. 65
MAGNPPPEGLGDDFFEQILAGQPPGYGGEAVGSTSLPMMGLQLGSSAANVGLMRSSANNNMGMMPL
GLNLEHHGFLRQQQDDGSSSLDTNNSNNINNASPFSITSAGFLGRDSVHMTSLFPTFGQLQIHSSI
RSAPPPLGPPQIHQFNSQPTSGAVSAVPQPPGIRPRVRARRGQATDPHSIAERLRRVRITERVKAL
QELVPTCNKTDRAAMLDEIVDYVKFLRLQVKVLSMSRLGAAGAVAQLVADVPLSSVQGEGIEGGAN
QQAWENWSNDGTEQEVAKLMEEDVGAAMQLLQSKALCIMPVSLASAIFRARPPNAPTLVKPESNPP
S

SEQ ID NO: 306, P. trichocarpa scaff XIII. 403
MAGNPPPEGLGDDFLEQILAAQPPGYGGGGEVMGSTSLPMMGLQLGSCAGNVGLLRSNANNNMGMM
PLGLNLEHHGFLRQQQDDSGSSLDTNNSNNINNTPPSIKSARILGRDSVHMTSLFPTFGHLQIHSS
IRPTPPPPGPPQIHQINSHPNPGAVSAVPQPPGIRPRVRARRGQATDPHSIAERLRRVRITERVKA
LQELVPTCNKTDRAAMLDEIVDYVKFLRLQIKVLSMSRLGAAGAVAQLVADVPLSSVQIKGEGNEG
GANQQSWENWSNDDTEQEVAKLMEEDVGAAMQFLQSKALCIMPISLASAIFRARPPNASTLINTES
NTPS

SEQ ID NO: 307, P. trichocarpa scaff28. 86
MISSNQSVESLATQDTSSVVLGSESDYAVDKVLISEQARLQNDCQNCNGNPSPDGMARGNLKFGNT
GLQCNGILPTLSSLNYPNQLPIVGDLTSYLSFSEASNAGCNGREQSEYLRSLKNLQNLSSIPQLWP
SQSYEGVSSLPPLMGQDRIEGSGLRGGNLDDDMHIMGKGYMGMDEILRLDKLSASPTTEGKEDLQS
CPFSSGIAEPNVNMSMNQLSSMPQTTSAAPVEGCNGTGKTRVRARRGHATDPHSIAERLRREKIAE
RMKNLQELVPNSNKVDKASMLDEIIEYVKFLQLQVKVLSMSRLGAAGAVIPLLTDGQPEGHNSLSL
SPSAGLGIDISPSADQIAFEQEVLKLLESDVTMAMQYLQSKGLCLMPIALAAAISSVKASLSGTTS
EERKNNGYTSGLVSSSSSITGIDTHPMSNDNNIATGTLSSKGMIVNGCNEVVKQEVLKNT

SEQ ID NO: 308, P. trichocarpa TC63334
MANNPTEPPTDDFLQEILGMPNFASAEAGLVGADAGLAGAAAAQASMMLQLSSGDGSGHISDLGGA
PGGGSAGFHGFPLGLSLEQGKGGFLKPEEASGSGKRFRDEIVDGRAKNVFHGQPMPTTVAIAPHPP
AMRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVTKTDRATMLDEIVDYVKFLRLQVK
VLSMSRLGGAGAVAPLVTDIPLSSVEDETGEGGRNQPAWEKWSNDGTERQVAKLMEENVGAAMQFL
QSKALCIMPISLATAIYHTQPPDTTTIVKPETNPPS

FIGURA 24 CONT.

SEQ ID NO: 309, P. trifoliata TA5414
MQPCSSGGEMQGISSLLSNGSHEQQSQIHQSATFDPTSQDDFLEQMLSSLPSCSWTDLKSPWGVVD
LNPNNNNNNINNKQPRDLSDETAPSTTQENNVPAGFQFDESMILASKMRQHQISGNTGSGNNNSAA
AKFMMQQQQQQQIMMAAARGGIGLPLSLGNGGDNDIVDASSFKSQQGGDGSVQALYNGFTGSLHGS
TQPQHFHHLQGGSMPGQTFGAPGPVMNQTQAQASGSTGGGGNTPAQQPKQRVRARRGQATDPHSIA
ERLRRERIAERMKALQELVPNANKTDKASMLDEIIDYVKFLQLQVKVLSMSRLGGAAAVAPLVADM
SSEGAGGGDCIQANGRNPNGAQTTSANDSLTVTEHQVAKLMEEDMGSAMQYLQGKGLCLMPISLAT
AISSATCHSRNPIISTSNNNNNNGNPHHNPLFQSNGEGPTSPSMSVLTVQSATMGNGGADGSVKDA
ASVSKP

SEQ ID NO: 310, R. communis TA1616
MAGNPPPEGLGDDFFEQILAVQPGYGGGDGGGGGDVVGSTMPMMGLQLGSASSGGGGLRTNNMGMM
PLGLNLEFLRQQEDGSGALDNNNHTNNNNNNVSSSTTSVINDRDSVHMASLFPTFGQLQTQTLRPP
PPPHLHQPFHGQPTPGVVSAASQPPAIRPRVRARRGQATDPHSIAERLRRERIAERMKALQELVPT
ANKTDRAAMIDEIVDYVKFLRLQVKVLSMSRLGAAGAVAQLVADVPLASVEGESIDGAAANQQTWE
KWSNDGTEQQVAKLMEEDIGAAMQFLQSKALCIMPISLASAILRTHPPDAPSIIKPESNTPS

SEQ ID NO: 311, R. communis TA2825
MANNPTEPPADDFLQEILGLPNFASADAAGLVGADGALATPMMLQLSSGDGSNHITALGGGGGGGG
GAGFHGFPLGLSLEQGKGGFLKPEEASGSGKRFRDDVVDGRANTVKNVFHGQPMPTTMAAAPHPPT
MRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVNKTDRAAMLDEIVDYVKFLRLQVKV
LSMSRLGGAGAVAPLVTDIPLSSVEDETGEGGRNQPAWEKWSNDGTERQVAKLMEENVGAAMQFLQ
SKALCIMPISLATAIYHTQAPDTSTIVKPETNPPS

SEQ ID NO: 312, S. bicolor TC104646
MAGQPPPPGGSEDDFLEHFFAFPSAASAAAGGHAGAGAGGDHPFPLALSLDAAAEPKPDRDPVQLA
GLFPPVFAGAGGVQQPHLRGPPPPQMFQAQPKPGEGGMAPQPPAPRPKVRARRGQATDPHSIAERL
RRERIAERMRALQELVPNTNKTDRAAMLDEILDYVKFLRLQVKVLSMSRLGGAGAVAQLVADIPLS
VKGEASDSGSTQHIWEKWSTDGTEKQVAKLMEEDIGAAMQFLQSKALCMMPISLAMAIYDTQHSQD
GHSLMKPEPNTSS

SEQ ID NO: 313, S. lycopersicum AK247217
MANNPSEGPSDDFFDQILGFPAYNGAEPNLAGNDAGAIPPAMMLQLNSGDGSSQFTGVGLGVGLGG
GGFHGHGGGGSFPLGLSLEQGKGGFLKMDDVSAPGRRFRDDVVDSRASSSVKPGFHGQPMPSMPHP
PAIRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVNKTDRAVMLDEIVDYIKFLRLQV
KVLSMSRLGGAGAVAPLVTDIPISSVEEESSEGGNNNQPAWEKWSSDGTERQVAKLMEENVGAAMQ
FLQSKALCIMPISLASAIYHSQPPDTSSLVKPETNPPS

SEQ ID NO: 314, S. lycopersicum TA21665
MADNPPEVYAADDFLEQILAIPSYASLPVTDLTAGASSENSTSGVSQLQQQPLFPLGLSLDNGFAD
ANNTGGFQVKTEREAMNMGNLYPGLEHLQSHAVCLSVPQVHQVQPFQGHPTSSAIVTIPHQPAIRP
RVRAQRGQATDPHSIAERLRRERISERIKALQELVPSCNKTDRAAMLDEILDYVKFLRLQVKVLSM
SRLGGTSAAAQVVADIPLQSVEGDTCESHSNQHVWEKWSDSETEQEVAKLMEEDVGTAMQYLQSKS
LCIMPISLAALIYPTQQSDNQSMVKPEQAAPL

FIGURA 24 CONT.

SEQ ID NO: 315, S. lycopersicum TC172581
MAANQPEGYADDFLEQILAIPPYSGLPVADVGTPSETTSFTSASAVSHLNSAAAAGLQQPLFPLGL
SLDNGRDDVGDAGPYAVKHERDGMNIGNLYAGLEHLQSHAVRHSVPSVHHVQPFQGPPTTSTTVTV
PHPPSIRPRVRARRGQATDPHSIAERLRRERISERIKALQELVPSCNKTDRAAMLDEILDYVKFLR
LQVKVLSMSRLGGASAVAQLVADIPLNLWEGDSGESRSNQHIWDKWSNVDTEREVAKLMEEDVGAA
MQYLQSKSLCIMPISLAALIYPTQQPDDQSLVKPEAAAPS

SEQ ID NO: 316, S. tuberosum TA30646
MANNPSEGPSDDFFDQIMGFPAYNGAETNLAGNDAGAIPPAMMLQLNSGDGSGQFTGVGLGVGLGG
GGFHGHGGGASFPLGLSLEQGKGGFLKMDDVSAPGRRFRDDVVDSRASSSVKPGFHGQPMPSMPHP
PAIRPRVRARRGQATDPHSIAERLRRERIAERIRALQELVPSVNKTDRAVMLDEIVDYVKFLRLQV
KVLSMSRLGGAGAVAPLVTDIPISSVEEESSEGGNNNQPAWEKWSSDGTERQVAKLMEENVGAAMQ
FLQSKALCIMPISLASAIYHSQPPDTSSLVKPETNPPS

SEQ ID NO: 317, S. tuberosum TA28621
MAANQPEAYADDFLEQILAIPSYSGLPVADVGTPSETTSFTSASAVSHLNSAAAAGLQQPLFPLGL
SLDNGRDDVSEAGAYAVKHERDGMNIGNLYAGLEHLQSHAVRHSVPSIHHVQPFQGPPTTSTTVTV
PHPPSIRPRVRARRGQATDPHSIAERLRRERISERIKALQELVPSCNKTDRAAMLDEILDYVKFLR
LQVKVLSMSRLGGASAVAQLVADIPLQSVEGDSGESRSNQHIWDKWSNVDTEQEVAKLMEEDVGAA
MQYLQSKSLCIMPISLAALIYPTQQPDDQSLVKPEAAAPS

SEQ ID NO: 318, S. tuberosum TA34455
MQAMNSLLSQQQQSQMSLQDLQNGGNGGSTGGVGGLGQHNMGHFHFDPTSSHDDFLEQILSSVPSS
SPWPDLSKSWDPHHHLSSPPHNPSSGEDQPPSNPLHSQFHYDDQASSLLASKLRQHQITGGGAAAA
AKALMLQQQLLLSRTLAGNGLRSPNGASGDNGLLSLPLNLSNNGDQNDGVANPANDNSVQALFNGF
TGSLGQTSNQPQHFHHPQGGSMQSQSFGAPAMNQTPAASGSAGGGGGSTPAAQPKQQRVRARRGQA
TDPHSIAERLRRERIAERMKALQELVPNANKTDKASMLDEIIDYVKFLQLQVKVLSMSRLGGAAAV
APLVADMSSEGRGEGNVGRGGNGTAASSSNKETMTVTEHQVAKLMEEDMGSAMQYLQGKGLCLMPI
SLATAISTSTRIPLLSPEAGRPASPTLSALTVQSATIGNAAGKDATSLFEA

SEQ ID NO: 319, S. tuberosum TA37666
MQAMNSFQSTGENGASSGEHISHSHFDPSSSHDDFLQQILSSVPSSSPWPEISGDGHPYNFDDHQS
TLLASKLRQHQINGGSSAAAAAKALMLQQQLLLSRGIAGNGGDQNDDGLNPGNDISVQALYNGFAG
SLGQTSNQSQHFHHSQAQSFGAPAASLTMNQTPAASGSAGGAQPKQQKVRARRGQATDPHSIAERL
RRERIAERMKSLQELVPNANKTDKASMLDEIIDYVRFLQLQVKVLSMSRLGGAAAVAPLVADRSSE
GGGDCVQGNGGRGGSNGTTSSANNDSSMTMTEHQVAKLMEEDMGSAMQYLQGKGLSLMPISLATAI
STSTCHSMKPNNPLLLGGGSAINGGGETGGGPSSPTLSASTVQSATMGNGGT

SEQ ID NO: 320, T. aestivum TC253044
MLLCSTIQVIQMGNLSVFDEATMASLHDSKEFLSGSISSFGTAEQSQLAGSGLKAEQQEQNAMCNI
PLPFASGSQMAVSEAQGAMIPSKISSTIHNNKSEYPVPISHSADAQNKANSANGNSASAKPRARAR
RGQATDPHSIAERLRREKISERMKNLQDLVPNSNKADKSSMLDEIIDYVKFLQLQVKVLSMSRLGA
PGAVLPLLAESQTEGRSNSPLSSPTASQGLLDAAGPEDSLVFEQEVIKLMETSITNAMQYLQNKGL
CLMPIALASAISNQKGTSAAAIPPER

FIGURA 24 CONT.

SEQ ID NO: 321, V. vinifera GSVIVT00016367001
MDEYLDHLFSSTSWSDVNVKEGSSWICGEPSQTNGMLSDSIGIYEGDKKNSPVGMTNSNLMIEGLV
TQDTSSIVLGGESDHGLGKGLLLEEAPRQQDLQNTEGNSSLNGAVNGSSEVGYVGLQLNTPTSTPC
SLDLGSPKQLSLIGGMSRSSRPFTELDHVGCNGNEPSDFQRSVGNFQALPPIPQLWSQPSYGGGSS
LSPVMGEYKMQGFGLQGEYVDNEMDIMRNRYVGDEILQLDNISSAIPIKGKEEQHTHPFSPSAVGP
HMTMTASGLQSLPQTTVGTASGGCNGTGKPRVRARRGQATDPHSIAERLRREKIAERMKNLQELVP
NSNKTDKASMLDEIIEYVKFLQLQVKVLSMSRLGAAEAVVPLITDGQAEGSKGLSLSPSAGQAEDI
CQSPDQIAFEQEVVKLMESNVTMAMQYLQSKGLCLMPIALATAISSGKAASSGTGSDEGKNAATPV
AATPVATAYLALGHIIHLLMAMF

SEQ ID NO: 322, V. vinifera GSVIVT00017237001
MLSTLPSWSDLPANPKSPWELNASNPISMPSNKSRDLSDDTTPSNPDNVQFAFDESAMLASKLRQH
QISGNSSAAKSALMLQQQLLLSRGVAMGRSPSNGSGAGESGLLQLPLSLSNGDSCLVDRSQNDVVD
GSSSFKSPNQGGDGSVQALYNGFAGALHGSGQASNQAQNFHHPQGGSMQAQNYGAPATRVRARRGQ
ATDPHSIAERLRRERIAERMKALQELVPNANKTDKASMLDEIIDYVKFLQLQVKVLSMSRLGGAAA
VAPLVADMSSEGGGDCIQASGTSGPTGGRATNGTQTTTSNDSLTVTEHQVAKLMEEDMGSAMQYLQ
GKGLCLMPISLATAISTTTCHSRNPMVAAAAVAASNINNGSHTHPLLPNSNADGPSSPSMSVLTVQ
SATMGNGLADAPVKDAASVSKP

SEQ ID NO: 323, V. vinifera TA46156
MASNPSEAPADDFLEQILGIPTYPAADPNLAANDVNLAAPMVLQLGSGEGSGHIAGGYQGTMFPLG
LRLEQGKSSFLKPEDASGSGKRFREEVIDGRASTVKNAFHGQPMQATVAAAPHPPAIRPRVRARRG
QATDPHSIAERLRRERIAERIRALQELVPSVNKTDRAAMLDEIVDYVKFLRLQVKVLSMSRLGGAG
AVAPLVTDIPLASVEEEASEGGRNEPAWEKWSNDGTERQVAKLMEENVGAAMQFLQSKALCIMPIS
LASAIYHSQQPDTTPLIKPQTNPPS

SEQ ID NO: 324, V. vinifera TA46194
MATNPPEGYADDFLEQILAIPSYPGHDPNMVGTGSTMVLQLSSGDGSGHVAGGGFQGPVFPLGLSL
ESGIPATQVAPGSGERFRDDVDARGSSGRSERESVHLDGLFPAYGHVQSLSVRPAVPQVHQAFHGQ
PTPVPITAAPHPPAIRPKVRARRGQATDPHSIAERLRRERIAERMKALQELVPSSNKTDRAAMLDE
IVDYVKFLRLQVKVLSMSRLGGAGAVAQLVADIPLPAVEGETGEGGSNQQAWDKWSNDGTEREVAK
LMEEDVGAAMQFLQSKALCIMPISLAAAIYPAHQTDTPTLIKPEPNAPS

SEQ ID NO: 325, Z. mays TA139285
MAGQPPPQGPEDDFFDQFFSMTAGGSYPGATAGGGRAPGDQPFSLALSLDAAAAEASGSGKHADGG
KADREAIQLPGLFPPAFGGGVQPPHLRATPPTQVFHAQQPKQGGAAVGPQPPAPRPKVRARRGQAT
DPHSIAERLRRERIAERMRALQELVPNTNKTDRAAMLDEILDYVKFLRLQVKVLSMSRLGGAGAVA
QLVADIPLSVKGEASDSGSKQQIWEKWSTDGTERQVAKLMEEDIGAAMQFLQSKALCMMPISLAMA
IYDTQHSQDGQPVKPEPNTPS

SEQ ID NO: 326, Z. mays TA126400
MAGQPPPSGASEDDFLEHFFAFPSAASAGAAGGHAGAGVGGDHPFPLALSLDAAAEAKPDRDPVQL
AGLFPPVFAGAGGVHQPHLRGPPPPQMFQAQPKPGEGGMAPQPPAPRPKVRARRGQATDPHSIAER
LRRERIAERMRALQELVPNTNKTDRAAMLDEILDYVKFLRLQVKVLSMSRLGGAGAVAQLVADIPL
SVKGEAGDGGGAPQQQQQQHVWEKWSTDGTEKQVAKLMEEDIGAAMQFLQSKALCMMPVSLAMAIY
DTQHPLDGHGHSLKPEPNASS

FIGURA 24 CONT.

SEQ ID NO: 327, A. cepa TC3698

CACTTCACTTCCCCTTTCTTCATTTCTTCAATACCTACAAACCCTAGCTTCACTTTACTATTTAAA
AATGGCGTCTAATAACAACCCGCAACCTCCTTCCGATGGACTCAACGACGACTTCTTCGACCAAAT
TTTTTCCATGCCTTCCGCCTTCGCCGCCTCCTCAGCCTCTACTGACGCCACCGAAGCTCTCAACTA
TGCTGATTCCTCCACCACTGGGGTTACCCCGATGTTCTCACTTGGCCTCAGCTTGGATCAGCAGCC
TAAGTCTGATTCGTCCCAATCGGAAAGAGAGCATCCGGTTCAGTTTGCTTGCTTGTTTCCACAGTA
TGGGCACAATAATCAAGTTCATCAAACCCGGCCTAATCTTTCTCCTCCTCAAGTTCTTCATGGACA
AGCAATGCAGAGCAGTATGGCTGCAACACCACAACCTACAGGTCCACGGCCAAGGGTTAGGGCTAG
GAGAGGCCAAGCTACCGATCCTCATAGTATAGCTGAGCGATTGAGGAGGGAAAGAATAGCAGAAAG
GATTAGAGCTTTGCAGGAATTGGTGCCCAGCACTAATAAGACCGATAGAGCTGTGATGCTGGATGA
AATTGTAGAGTATGTGAAATTCTTGAGGCTTCAAGTAAAGGTACTAAGCATGAGTAGACTAGGTGG
TGCTGGTGCTGTTGCACAGCTGGTAGCCGATATCCCTATGTCTTCTTCAGTTGAGGGTGATTCAAG
TGAAAGTGGGAAGACAAGCTACCAAGGGTGGGAAAAATGGTCAACCGACGGCACCGAGCGTCAGGT
GGCAAAACTAATGGAAGAAGATGTAGGGGCTGCAATGCAATTTCTCCAATCAAAAGCCCTCTGCAT
CATGCCCATATCTTTAGCAGCTGCAATTTGTCAAACGAATCCCTCAACCGAAATCTCTTTCATG

SEQ ID NO: 328, A. thaliana bHLH007-11-AT1G03040

GTCTGTGTCTCTTCGTCCTCATTAGCTTAAGCATCATCATTCTCTCACCTAACAACTCCACACAAG
ATTCTTCGCATGGAAAGTTGTTCTTCGACTTCTCTTCTCTAACTCGCTATCTTTTAACTCACCCAG
CTCCACTGAGTCGAAAATTTCAAACCTTTACTCGTTTCCTTCATGGCTAATAACAACAACATCCCA
CATGATAGCATCTCCGATCCATCTCCTACCGACGATTTCTTCGAGCAGATCCTCGGGCTTTCCAAC
TTCTCCGGTTCTTCAGGTTCTGGTCTCTCTGGAATCGGCGGCGTGGGTCCACCTCCGATGATGCTT
CAGCTTGGTTCAGGCAACGAAGGGAATCATAATCATATGGGTGCCATTGGAGGAGGTGGACCTGTA
GGGTTTCATAATCAGATGTTTCCGTTGGGATTAAGTCTCGATCAAGGGAAAGGACATGGCTTTCTT
AAACCTGATGAAACTGGTAAACGTTTCCAAGACGATGTTCTTGATAATCGATGTTCCTCTATGAAA
CCTATTTTCCATGGGCAGCCAATGTCACAGCCAGCTCCACCAATGCCGCATCAACAGTCTACTATT
CGGCCTAGAGTTAGGGCTAGGCGAGGTCAAGCTACCGATCCACATAGCATCGCTGAGAGGCTCCGA
AGGGAAAGAATAGCAGAACGGATCAGGTCGTTGCAGGAACTTGTACCTACCGTTAACAAGACAGAT
AGGGCTGCTATGATCGACGAGATTGTCGATTATGTAAAGTTTCTCAGGCTCCAAGTTAAGGTCCTG
AGCATGAGCCGTCTTGGTGGAGCCGGTGCTGTCGCACCACTAGTCACTGAAATGCCATTATCTTCA
TCAGTTGAGGATGAGACGCAGGCCGTGTGGGAGAAATGGTCAAACGATGGGACAGAGAGGCAAGTG
GCTAAGCTGATGGAAGAAAACGTTGGAGCAGCGATGCAACTTTTGCAATCAAAGGCTCTTTGCATA
ATGCCGATCTCATTGGCAATGGCGATTTACCATTCTCAGCCACCAGACACATCTTCTTCAATCGTC
AAACCAGAGATGAATCCTCCACCGTAGATTTTTGTTCATCCAACGGTCCCCAGCTGATGATTGACA
TTTTGCTCTGTTTCCCACTACTAGACTTTTGTGACTCATGAAAGGTAAGTAAAAAGGCATTGGAGA
TGGAATCTAAGTAGGATTTGTGCAGTAAAGAAGTAAAACGGGATCTGTCAAAAGAAGGAAAAAGCT
CTCGCTTGCTTGGCTAGTATTTATCATTTTGATGAAAGTAACTCTTTTTTGTTCAAAGACTTTAGT
GTGATTTTCAGGACCAAGGGCTTTGAGGGTAGTGCTAGCTGTAGTAATAGTAATGAAGGTGTGGGA
TCGTGTTTCTGAATTATGTAAAAAAGGAAGAAAAAACAAATGTTGGTATTATATTATGGTTTTGCC
TCTC

SEQ ID NO: 329, A. thaliana bHLH059-11-AT4G02590

GCTCCTTTCTCGTCTCTGTCTTCTTCGTCCTCATTCGTTTTAAAGCATCAAAATTTCATCAACCCA
AAATAGATTAAAAAAATCTGTAGCTTTCGCATGTAAATCTCTCTTTGAAGGTTCCTAACTCGTTAA
TCGTAACTCACAGTGACTCGTTCGAGTCAAAGTCTCTGTCTTTAGCTCAAACCATGGCTAGTAACA
ACCCTCACGACAACCTTTCTGACCAAACTCCTTCTGATGATTTCTTCGAGCAAATCCTCGGCCTTC
CTAACTTCTCAGCCTCTTCTGCCGCCGGTTTATCTGGAGTTGACGGAGGATTAGGTGGTGGAGCAC
CGCCTATGATGCTGCAGTTGGGTTCCGGAGAAGAAGGAAGTCACATGGGTGGCTTAGGAGGAAGTG

FIGURA 24 CONT.

GACCAACTGGGTTTCACAATCAGATGTTTCCTTTGGGGTTAAGTCTTGATCAAGGGAAAGGACCTG
GGTTTCTTAGACCTGAAGGAGGACATGGAAGTGGGAAAAGATTCTCAGATGATGTTGTTGATAATC
GATGTTCTTCTATGAAACCTGTTTTCCACGGGCAGCCTATGCAACAGCCACCTCCATCGGCCCCAC
ATCAGCCTACTTCAATCCGTCCCAGGGTTCGAGCTAGGCGTGGTCAGGCTACTGATCCACATAGCA
TCGCTGAGCGGCTACGTAGAGAAAGAATAGCAGAACGGATCAGGGCGCTGCAGGAACTTGTACCTA
CTGTGAACAAGACCGATAGAGCTGCTATGATCGATGAGATTGTCGATTATGTAAAGTTTCTCAGGC
TCCAAGTCAAGGTTTTGAGCATGAGCCGACTTGGTGGAGCCGGTGCGGTTGCTCCACTTGTTACTG
ATATGCCTCTTTCATCATCAGTTGAGGATGAAACGGGTGAGGGTGGAAGGACTCCGCAACCAGCGT
GGGAGAAATGGTCTAACGATGGGACTGAACGTCAAGTGGCTAAACTGATGGAAGAGAACGTTGGAG
CCGCGATGCAGCTTCTTCAATCAAAGGCTCTTTGTATGATGCCAATCTCATTGGCAATGGCAATTT
ACCATTCTCAACCTCCGGATACATCTTCAGTGGTCAAGCCTGAGAACAATCCTCCACAGTAGGATT
TCTGCAATAAAGAGTTTGTACAGCTAATCCAACTGTCCAACATGGGTTTTTCTTCTGCTCTAATGA
CTCTGGTTTCTTCTCTCCTCTCTCACCCACTTGAAAGGTAAAAAAGTGAAAAAGGCTTTGTAGATG
GAATCAATGTAGGATTTGCAGTAGAGGGAAAAAAAATGTCAAAAAGCTCAATTGATCAAGTATTAT
TGTAATCATTGTACCTTTATTTTAGGTGGACTTTGATGAAAGCAACTTTTTGTTTTCAAGACTTTA
GTGGGAGGTTGAGGAAGGAGCTTGAAGGGTGTTATTTATTAGTAGTAGTAGTAGTGGGAAGTTGTG
GGACCTTGTTGAGTTGTGTTCAAATTGAAGAAAAAACAAGTATTTGTAATTTGTCACCCCTTGTAT
TATTATTTATTTTGTATGACTTTGGAGTAGTATAATTAATTATCAAATAATATTATTAGTTTGATC
TAGT

SEQ ID NO: 330, A. thaliana bHLH066-11-AT2G24260

TCTTCCCAAAAAAAAGTGTAGAAGCAGAAGAAACCCATCATCATCATGATGAACTCTTCTCTTCTA
ACTCCTTCTTCTTCATCTTCTTCCCATATCCAAACTCCATCAACAACTTTCGACCACGAAGACTTC
CTCGATCAAATCTTTTCCTCGGCCCCGTGGCCCTCCGTCGTCGATGATGCTCATCCTCTTCCCTCC
GATGGCTTCCACGGCCACGATGTCGACTCAAGGAATCAGCCGATCATGATGATGCCTTTGAATGAT
GGCTCCTCCGTCCACGCTCTTTATAATGGCTTCTCCGTCGCCGGATCTCTTCCTAACTTTCAAATC
CCTCAGGGATCGGGAGGAGGATTGATGAACCAACAAGGACAAACGCAAACGCAAACGCAACCTCAG
GCGAGCGCGTCTACAGCTACTGGTGGTACGGTGGCGGCTCCGCCGCAGAGTAGGACTAAAATCCGA
GCTAGGAGAGGTCAAGCAACTGATCCTCATAGTATCGCCGAAAGGTTACGAAGAGAGAGAATTGCG
GAAAGAATGAAAGCTCTTCAAGAACTCGTTCCTAACGGCAATAAGACAGACAAGGCATCGATGCTC
GATGAGATCATAGATTATGTCAAGTTTTTACAACTCCAAGTCAAGGTACTGAGCATGAGCAGATTG
GGAGGTGCTGCTTCCGTTTCTTCTCAAATCTCCGAGGCTGGTGGATCCCACGGGAACGCATCCTCC
GCCATGGTCGGCGGTAGCCAGACGGCCGGAAACTCCAACGACAGCGTTACAATGACGGAACATCAA
GTGGCCAAACTAATGGAAGAAGACATGGGCTCGGCCATGCAATATCTTCAAGGGAAAGGTCTTTGT
CTCATGCCAATCTCTTTAGCCACCGCCATCTCAACCGCCACGTGTCACTCCCGTAACCCTTTGATC
CCTGGAGCTGTTGCCGACGTCGGAGGTCCTTCCCCTCCCAATCTTTCTGGCATGACCATACAGTCG
ACGAGTACAAAAATGGGTAGCGGTAATGGGAAATTAAACGGTAACGGCGTGACCGAGAGGTCGTCT
TCTATCGCCGTTAAAGAGGCCGTATCCGTTTCGAAAGCGTGATAACGGCCGTTTACTTTTGGGGTT
AGGCAGAGAGATACCAAAAACAAAGAGAAAGTGGGGTGAAGTGGGAGATAAAGTCAAAGTCTACGA
AGAATGAGGTTGAGGTTGTTTGATGTGTTCTCTGAACGAAGCCATATTTTTCGTAGAGTTTTGGTT
CTTTTACCTACGCACTACAAAACCATTCAAGGGACATTTCTTTTTCTTTTTTAAATATATGGTTGC
AATATGTTTTATTTTTATTAAATAAAGTTGGTATCGTTGATCTGAATATTAGGCTAGAAAAAGAAA
AAAATAGTATATTGGTTTTTC

SEQ ID NO: 331, A. thaliana bHLH069-11-AT4G30980

AACATTCATTTTCATTCTCACTCCCCACCAAAAAAAAAAAAAAACAGAAGCCATGAACTCCTCGTCT
CTTCTAACTCCTTCATCATCTCCTTCTCCACATCTTCAATCTCCTGCAACATTCGACCACGATGAT
TTCCTCCACCACATCTTCTCCTCCACTCCTTGGCCCTCATCCGTTCTCGACGACACTCCTCCACCA

FIGURA 24 CONT.

ACTTCCGATTGTGCCCCCGTCACTGGATTCCACCACCACGACGCCGATTCAAGAAACCAGATCACT
ATGATTCCTTTGTCACATAACCATCCTAATGACGCTCTCTTCAATGGCTTCTCCACCGGATCTCTC
CCTTTCCACCTCCCTCAAGGATCGGGAGGTCAAACGCAAACGCAGTCGCAGGCGACGGCGTCAGCC
ACCACCGGTGGTGCAACGGCGCAACCTCAGACAAAGCCTAAAGTCCGAGCTAGGAGAGGTCAAGCC
ACTGATCCTCACAGTATCGCCGAACGGTTACGGAGAGAGAGGATAGCGGAAAGAATGAAATCTCTT
CAAGAACTTGTCCCTAATGGTAACAAGACAGACAAAGCATCAATGCTCGATGAGATTATCGATTAT
GTCAAGTTCTTACAGCTCCAAGTCAAGGTACTAAGCATGAGTAGACTGGGCGGTGCTGCTTCTGCT
TCTTCTCAAATCTCTGAGGATGCCGGTGGATCCCACGAAAACACCTCCTCCTCCGGCGAGGCGAAG
ATGACGGAGCACCAAGTTGCAAAGCTAATGGAAGAGGACATGGGATCAGCCATGCAATATCTACAA
GGCAAAGGTCTTTGCCTCATGCCCATCTCGTTAGCCACCACCATCTCCACCGCCACGTGTCCTTCT
CGTAGCCCCTTCGTTAAAGATACCGGCGTTCCTTTGTCTCCTAACCTATCCACTACAATAGTTGCT
AACGGTAATGGCTCATCGTTGGTCACCGTTAAAGACGCTCCCTCCGTTTCCAAGCCGTGATAACGG
CCATTTGTCCATTTCATTTTCCCTTTTTTGGGTGGGAAAGAGAGAAAAAAGTTTAGAAGACAAAGA
CAAGTGGGATAGGTGGTTTTGGTCAAAGTTTAGAAAGAATAAGGTCGTGTTTTCGGATACGACACC
GTATTTGCGTACACTTTGGTTTTCTGTCTTTACCTACTACAAACCACCCATAAGCACACTCATGTT
ATCATGTTTTTTTTTTTTTGGTTTATAAAGTTATATCCTTAA

SEQ ID NO: 332, A. thaliana bHLH082-11-AT5G58010

ATGGAAAATGGAAATGGAGAAGGAAAAGGAGAATTCATAAACCAAAACAATGACTTCTTCCTTGAT
TCCATGTCAATGCTCTCCTCTCTCCCTCCTTGTTGGGATCCATCTCTTCCTCCTCCTCCTCCTCCG
CCACAATCTCTTTTCCACGCTCTAGCCGTCGACGCCCCCTTCCCCGACCAGTTCCATCATCCTCAG
GAGTCAGGAGGTCCAACAATGGGTAGCCAAGAAGGGTTACAGCCACAAGGTACAGTGTCGACCACG
AGCGCACCTGTTGTTCGTCAAAAGCCAAGAGTTCGAGCCAGGAGAGGCCAAGCCACCGATCCTCAC
AGTATCGCTGAGAGGCTTAGAAGGGAACGCATCGCAGAGCGTATGAAGTCTCTCCAAGAACTGGTT
CCCAACACCAACAAGACGGATAAGGCATCAATGTTGGACGAGATCATCGAGTATGTTAGGTTCCTT
CAGCTTCAAGTCAAAGTACTAAGCATGAGCAGATTGGGAGGTGCAGGATCAGTTGGTCCACGCCTC
AATGGTCTCTCCGCAGAGGCAGGAGGACGGCTCAATGCCCTCACTGCACCGTGCAATGGCTTAAAT
GGGAATGGAAACGCTACAGGATCGTCCAATGAGAGCCTAAGGTCAACAGAACAAAGAGTGGCGAAA
CTGATGGAGGAAGACATGGGATCAGCAATGCAATACCTTCAAGGGAAGGGGCTTTGCTTAATGCCA
ATATCTCTCGCAACCGCGATATCATCATCAACTACTCACTCTCGTGGATCCCTCTTTAACCCCATC
TCCAGTGCTGTAGCAGCAGAGGATTCAAATGTAACAGCAACTGCAGTTGCTGCACCTGAAGCCTCG
TCTACTATGGATGATGTATCTGCTTCCAAGGCCTGAACCATAGAATACTCATGTGTTTCGTTTTCT
CAGTTTAAGTCAAACGCTAATCTCCCTATCTTTTTACTATTTTCTCTGCTAAGCAACAATAATATA
AAACAGTTGCAACAAAGAACAAAACCGAC

SEQ ID NO: 333, A. vulgaris TC15501

CACTTTCAAAACTCTCTCTCTCTTTCTCTTCCCGAAACTCAAAATTTCTCCGGTTTATCGGGTATC
GGATCTGTAACTTGTTTCGCCGACAACCTCGTTTTTCGATTTTCTCTCATAAACCCAAAACAAAAC
CCTTTGTCTATTTCTATGTTTAATGTTCTTCAAAAACCCTAACACTTCAATCAAACACGCATACAC
TCTTACTCTCTTCTTTCTCACAACATGGCTAATAATAATAATCCTACAGAAGGATTAACAGATGAT
TTCCTTGATCAGATCTTATCTTATGGTGCTCACGAAGGTAATTTGGCGGGAAATGATGTTAATTTG
GCGGGAACAACTACACCTGGAGGTTCGATGGCCTTGCAGTTGAGTTCTTCTGGAGGTGATATTACC
GGTAGTGGTGGGTATCATCATCATCATCAGCATCAACATCATCATCAGTTAAGTGGTGGTGGTGGT
GGTTTTCCGTTAGGGTTAAGTTTAGAACAAAGTGGTAAAGGTGGTTTTTTTAAACCAGATGAAGCT
TCTGGAAGTGGTAAGCGGTTTCGTGACGATTTCGTTGATTTGAAAGCTGTTTCTTCTACTAACAAC
GATAACAACGATAGAGGAGACTCTGTACATTTGAATAATTTGTTTCCAGCATTTGGACATGTGCAA
CAACAGCAACAGGCTCACTCTGTTCGACCTCCTTCACATCAAAATCAAATCAACCAGCCTTTCCTT
GGGCATCCAACACCTGGGGCTGTTGCTGTTGTACCACATCCACCTGCCATTCGGCCTAGAGTGCGA

FIGURA 24 CONT.

GCAAGAAGGGGGCAAGCCACTGATCCTCACAGTATTGCAGAGAGGTTACGTAGAGAAAGAATAGCT
GAAAGAATGAAGGCACTACAGGAACTGGTTCCTAGTTCCAACAAGACTGATAGGGCAGCTATGCTT
GATGAAATTGTTGATTATGTTAAATTTCTAAGGCTCCAGGTCAAGGTTTTGAGCATGAGCAGACTG
GGAGGAGCTGGTGCAGTAGCACAGCTTGTGGCTGACATTCCACTTTCATCAGCTGAGGGCGGAGCC
AGTGAAGGTGGCAGTAATCAGCCAGCATGGGAGAAATGGTCTACAGATGGCACAGAACATCAAGTA
GCAAAGCTCATGGAAGAAGATGTTGGAGCTGCCATGCAATTCCTCCAATCTAAAGCACTTTGCATC
ATGCCCATTTCACTTGCCTCTGCAATATACCACGACCAACCACCCGATGCCTCCACAATGATCAAA
CCCGAACCTAATGCACCTTCATAAAATGAAAAGGATACATGCAATTGGACGTGCCCATGAACAACT
CCAAACTTACAACTATCCCACCTTCCATTTGTTCCCCCCACGTCAGACTTAGTCAACTCCCCTCAG
CATTGCAATCAAAGTCAGATGCCATATTATTCGTTCTGCCTTTCATTTTTTTCGTCAAATAGGTTG
TAATGTGATGCCTAGAAAAATAATCATAGGTACTCCTGATGTGGTGTGGGATAGTGGTGGATAGCC
AGAACACTTTCCGGTGGGGACCCTTCTTTTGGATAGTGGAGGGTGGGGCATTTTTAAAGGGGGAAG
ATGGGTACATAGACATGATTAATGTATTAGAAGGAAACTGAAAAGGGAATTGAAATGGATATTGGA
TAGTTGTTGGTCGCTTAAAAAAGTTGTGTTGGAAGGTTGCACTACTCGGTTTTGAGTGGGTTGTGT
TTTTTGTAAAGTGCTTGATTTATTTTTCTCAAGTATCTCTGAGTCAT

SEQ ID NO: 334, A. vulgaris TC19278

TCAACTGCTGTTGGCAGTTGTAATGGGGCTGTGAAGCTCGTGTGAGAGCTCGCAGGGGTCAAGCAA
CGGATCCACACAGTATTGCTGAAAGGCTTCGAAGAGAGAAAATAGCAGAGAGGATGAAGAATCTAC
AAGAACTTGTTCCAAACTCTAATAAGACGGATAAAGCATCTATGCTTGATGAAATCATTGAGTATG
TCAGGTTTCTTCAGCTTCAAGTCAAGGTTCTAAGCATGAGCAGGCTGGGGGCAACAGGGGCAGTTG
TTCCCCTCATTACAGAAAATCAGCCTGAGGGATCTACTAGTCATTTTCTGTCGCAGTCAGCTGGTC
GAGATGTTAAACATATTGAATCTCCTAACACCCTTACATTTGAGCAAGAAGTAGTGAAGCTGATGG
AATCCAACATGACCAAAGCTATGCAATTTCTGCAGAGAAAAGGCCTCTGCCTAATGCCTATTGCCC
TTGCAGCTGCCATTTCCGGTAGTAAAGCACCACCCGCCTCTCTTTCTACTGAGGGAAAGAAGTCTA
TTTTGACCAATGGCTTTGTTCATCATAACAGTGGGCAATCTAACATTGGGCCAGGTCAGATATCAT
CTGATGGAGAGACTAGATCAGAGGAAAACATAATTGGGATTCACAGCAGAGAAGACTTTTCAAGCA
ACAGCTGCAGTGGGACCAGCATTAAAAAAGAGGGAACAAACACCACAAACTTCACCAGAGAAATGA
ATCAAGAGTAGACATAGTTTTGCCGGTCTCAATCTTTGTGGCTTCATACCTGGTCTCCAATTCTCA
AGTTTCAGTAACCTAGCAAGTAGATAAGTAGCAGAGTCTGGGACTGTTTTGTTTTTTTTTTCTCGC
TTCAAAGATGAGAATGTGGTATAGATTTCTAATTATACGTTATAGGTTTCATTGTACATGCACACA
TATATGC

SEQ ID NO: 335, B. napus TC10015 part

CCCACGCGTCCGAAACACCTTCCGATGACTTCTTCGAGCAAATCCTCGGTCTCCCCAACTTCTCAG
CCTCTTCATCCGACGGCGGGTTAGGCGGAGGAGGAGCACCGCCGATGATGCTGCAGCTGGGCTCCG
GCGAGGAAGGAGGTCACATGGGTGGGCTAGGAGGAGGAAGTGGGTTTCACAACCAGATGTTTCCTC
TAGGTTTGAGTCTTGAACAAGGCAAAGGACAAGGCTTTCTTAGACCCGAAGGTGGTAGTCTTGGAA
CTGGGAAACGTTTCTCTGATGACGTCATGAAACCCGTTTTTCATGGGCAGCCAATGCAACAACAGC
CAGCTCCAGCGGCGCCGCATCAGCCTACATCGATCCGTCCCAGGGTTCGAGCTAGGCGTGGTCAGG
CTACTGACCCACATAGCATAGCTGAAAGGCTTCGTAGGGAAAGAATAGCGGAGCGGATCAGGGCGT
TGCAAGAGCTTGTACCTACTGTTAACAAGACAGATAGAGCTGCTATGATTGATGAGATTGTGGATT
ATGTAAAGTTTCTCAGGCTCCAAGTTAAGGTTTTGAGCATGAGTCGTCTTGGTGGAGCCGGTGCAG
GTGCTCCACTTGTTACTGATATGCCTTTGTCATCATCAGTTGAGGATGAATCTGGTGAAGGTGGAA
GGGCACCACAACCTGCGTGGGAGAAATGGTCTAACGATGGCACTGAACGCCAAGTCGCTAAACTGA
TGGAAGAAAACGTTGGAGCAGCGATGCAGCTTCTTTCATCTAAGGCGCTTTGTATGATGCCAATCT
CATTGGCTATGGCTATTTACCATTCTCAGCCTCCAGATACAACTTC

FIGURA 24 CONT.

SEQ ID NO: 336, C. clementina DY268946

CAACCGTGTTACACCGACCCATCTGCTGCTGCTTCAAACTCCGCTGATGACATACCCATGTTCCTA
CAGCTGGGCTCCGGCGGCCCTTCTCCTGGTGCCGGAGGAGGCGCCACCAGTGCTGCTGACATAAGA
AGTTTGGCTATGGGAATCAGCATGGGAATGATGCCTCTAGGGTTAAATTTGGAGCATAGCTTCTTA
AGGCAGCATGAAGATCATAATAGTAACCATAACAACAACAACAACACTAATGCTTCTTCTTCTTCT
CCTAATAATTCTTCTGCCACTTCTGGAATCAACGAGAGAGATTCCGTGCACATGCCAAGTTTGTTT
CCAGCGTTTGGACAGTTGCAAAGTCAATCAGCCCGGCCGCCGCTGCCACGTCCTCCTCAAGTACAG
CAGTTTCAGAGCCAACCAGCAGCCGCACCGCAGCCTCCAGCTGTCCGTCCAAGGGTGCGAGCAAGA
CGAGGGCAAGCCACGGATCCTCACAGCATTGCCGAGCGGTTGCGTAGGGAAAGAATTGCGGATAGA
ATGAGGGCTCTGCAAGAGCTGGTTCCTAGCTGCAACAAGACGGATAGGGCAGCCATGCTTGATGAA
ATCGTGGATTACGTGAAGTTTCTAAGGCTTCAGGTCAAGGTTTTGAGCATGAGTAGACTGGGTGCA
GCTGGTGCCGTGGCTCAGCTTGTAGCTGATGTTCCTCTGTCATCAGCTCTTGAGGGAGAGAGCATT
GATGGTGGAAGCAGTCAGCCAGAATGGGAGAAATGGTCAAATGACGGAACTGAACAACAAGTAGCC
AAGCTAATGGAAGAAGACATTGGAGCTGCTATGCAGTTCCTTCAATCCAAGGCACTTTGCATCATG
CCCATATCACTTGCCAGTGCAATCTATCGCACCCGTCAACCGGATGCCCCAGCATTCGTCAAGCCT
GAATCGAGCACCCCTTCATATTGTCTAAACAACAATTGGAGTCAGCAGTTGGCTCACAATGACTCA
AAATATTTAATCACGGGTACCTATTCCTATAGAATACTTTTGGTCCTCTTATCCAACCAAATAATT
CAAATTCCAATGCAACGTTCCCATTCTTCTTTCCCTTAATTAACCTAGGGCAAAAAAGGGAAATGA
AAAGAAGGCAATGGGCAATGGCAAGGGGGAAAAAGGGGAATTTTTTTCCAAAGTTTTTTTTTTTTA
TAAAAACTCCCCTTTTCAAAAAAAATTATAAAGGGGGGGGGGGGTTTTTTAGGGGGGGGGGAAAA
AACCCCCAAAAAAGGTTTTCGGGGGGTGTTTTATATG

SEQ ID NO: 337, C. longa TA2544

GCACCCGTTGAGACCTAACCTTCCCCGGGAAACTGGCGGCTGTGCCTATGCAACCGAGCAGCAAAG
AGACGGAAAGGATGGCGCAGTCCCACGCCGCCCTCCAGCTGGAGCTCCAGAACGGCGGCGGAGGCG
CCCAGAACGACGATTTCTTCGAGCAGACGATGTCCGGCTTCTCCACCGCCGCCTGGGCACCGGAGT
TGGGGACCCCCAGGTCGCTTTTCGGGGCGCGGTCTTCGGAGGAGTCGCCGGACGACGAGGGATTGA
ACTACGCCCCACCGTACGGTGGGTCGTCTCTTCTCGCCTCGCGGCACCGGATGCACCTCGGCACCT
CGCCGGCGGATTCGTCGATGTTTCTCCAGCTCGGTGGGCAGATTCTGTTACATCCTACGGCCGGCG
CCTCCGGGGGCGAGTCCGGGGGCTTCCTCCGGTTGCCCCTCTCCCTCGGTAGCGGAGGCTCTGGGG
TCTCGACTTCTGCCATCTTCGGCGACAGATCCCGAGGAGAAATCGACCCGCCGTTCGAATCGTCCA
ATCCCACCGAGGCTGAGGTGTTGTATAACGACGGATTCACCGGATCGCTTCCACGGGCGGTGCAAG
CGTCCCTTCATCAGCAACTCCTCCGACAACCCCAGAACTACGGAGCTGCTACGCTGACCGGACAGG
CTCTTGCGATAACAGCAACGGCTTCAGGTACCGCCGGCGCAGGTACTGCGCCACCCAAGCCGAGGG
TGAGGGCCAGAAGAGGTCAAGCGACCGATCCCCATAGCATTGCCGAAAGACTTCGGAGGGAGAGAA
TCGCAGAGCGTATGAAAGCTCTGCAGGAATTGGTGCCCAACGCTAACAAGACGGACAAGGCGTCGA
TGCTGGACGAGATCATCGACTACCTTAAATTTCTCCAGCTCCAAGTCAAGGTCTTGAGCATGAGCA
GATTGGGCGGAGCAGCCGCCGTCGCCCCGCTCGTCGCTGACATGTCTTCAGAGAGCGGCGGCGGAG
CGGGAACGCCGAGGAGTAATAACGATGGCCTGACGGCGGCGGAGCAGCAGGTGGCGAAGTTGCTGG
AGGAGGACATGGGTTCAGCGATGCAGTACCTGCAGGGGAAAGGCCTCTGCCTCATGCCCATCTCCC
TCGCCTCCGCCATCTCCTCCGCCACCTGCCGTCCTCGGCCTCCCCCCGGTGTCAACCTCCGCTACC
CGGCAGCCGGCGACGCGCCTCCGTCGCCGAGCATATCGGCGCTGACAGTCCAATCGTCCAACGCCG
GCAGCGCCGGCGCCGACGCCAGCTAGGATGGAGCTGTTCGTTAAGATCGCTGTCCAAAAAGTGCTT
GGAATTGGAATTTCAATCCTTAGT

FIGURA 24 CONT.

SEQ ID NO: 338, C. paradisi híbrido_TA3392

CGACTTTGCCGAGTCAAACTCATGGCGAACAACCCGAACGAAGCCTCCTCGACTGACGATTTCCTC
GAGCAAATCCTCGGAATTCCTAATTTCGGTTCTGCGGAGTCCGGCTTGGCGGCTAGCGATGGCGGC
TTGGCTGCAGGCTCGCCGATGATGCTGCAGCTGAGCTCCGGCGACGGCTCCAGCCACATCTCGGCT
CTCGGAGGCGGAGTAAGTAGTGGGTATCATGGGCAGGTGTTTCCGTTAGGCTTGAGCTTGGAGCAA
GGTAAAGGAGGATTTTTGAAGCCCGAGGAAGCCTCTGGAAGCGGAAAGCGGTTTCCTGAAGAACAT
GCTATAAAAAATGTTTTTCATGGCCAACCACTGCCCTCACCTGTTCCTGCAGCTCCACATCCACCA
GCAATGCGCCCCAGGGTGCGAGCTAGAAGAGGACAGGCCACTGATCCACACAGCATTGCTGAACGT
CTACGAAGAGAAAGAATAGCAGAAAGAATTAGGGCACTGCAGGAGCTGGTCCCTAGTGTCAACAAG
ACTGATCGAGCAGCCATGCTTGATGAAATTGTGGATTATGTGAAGTTTTTGAGGCTTCAAGTGAAG
GTTTTGAGTATGAGTAGAGTGGGCGCTCCTGGCGCTGTGGCTCCACTGGTAACAACCGACCTCCCA
CTATCATCAGTCGAGGATGAATCTGGTGAAGGGGTAAGAAACCAACCAGCCTGGGAGAAATGGTCA
AATGATGGCACAGAGCGACAGGTAGCTAAGCTCATGGAAGAAAATGTTGGGGCTGCCATGCAGTTC
CTTCAATCAAAGGCTCTTTGCATAATGCCGATCTCACTGGCCACGGCAATTTACCATTCGCAGCCA
CCGGAATGAAGAGGTAGGATGGATAATGGTTGGTTGTTTGAAAATGCTTAGTGTTTTCCCAGTCCA
ACATAAAATGATAGTGCTTGCTTTGTCGAGGAGGCGTGGGCTTTGTTTTCAAAAAGTAATTTGACT
TCTTTCTCATTTTCTCCTGAGTCATTCATTATTGGAGGGCGACGGGGCTGCTGACGATGGCTGATG
GCTGATGGCATTGGTATTGGTGGACTGCGTTTGGTATAAATTACTTACGTATTTGAACTTAACCTG
ATTTTCAATTCTAAAGTCATTTACTGCGGTACAAAATTTGAAAAATCTATTACTGAATTTAGTTGA
TAATAACCAACTTAGCGTGTTCCTTTCGGGGATCATCCATCCAACTC

SEQ ID NO: 339, C. sinesis TA12416

GGGGAAACACCAAATTCCTTCTCTCCCTCTCTCTCCCTCTCTCTCTCTCTTCTTCCTGTTCTCACC
TATTTAGTATTTAACATCCATCAGAAGAATGGTTTCCATTCTCTTGAAGAAACTCTTGACTCTTCA
AACTTAAATTTCCATATCGTAAAGACTTATTCTCAAGAAATTAAGATGCAGCCCTGCAGCAGCGGC
GGAGAAATGCAAGGAATCAGCTCGCTCTTGAGCAACGGAAGCCACGAGCAGCAGTCGCAGATCCAT
CAAAGCGCGACGTTTGATCCAACTTCGCAAGACGACTTCTTAGAGCAAATGCTATCCTCTCTCCCA
TCGTGTTCTTGGACTGACTTGAAGTCTCCCTGGGGGGTTGTTGACCTTAACCCTAATAACAATAAT
AATAATATTAACAACAAGCAGCCCAGAGACTTATCCGACGAAACGGCGCCGTCCACTACTCAAGAG
AATAATGTCCCTGCAGGGTTTCAGTTCGACGAGTCCATGATTTTAGCTTCCAAGATGAGACAGCAT
CAGATCAGCGGGAATACTGGGAGTGGGAATAATAATTCTGCTGCAGCAAAGTTTATGATGCAGCAG
CAGCAGCAACAAATCATGATGGCTGCTGCTAGAGGTGGCATCGGGTTGCCTTTATCACTTGGAAAT
GGTGGCGATAACGACATCGTTGATGTCTCGTCATTCAAGTCCCAACAGGGAGGAGATGGCTCAGTG
CAAGCACTCTACAATGGCTTCACTGGATCTTTGCATGGCTCAACCCAGCCTCAACATTTTCATCAT
CTTCAGGGAGGATCGATGCCCGGGCAGACCTTTGGGGCACCAGGGCCGGTTATGAACCAGACGCAG
GCTCAGGCAAGTGGGTCAACCGGTGGTGGTGGTGGTGGTGGAAACACACCCGCTCAGCAGCCTAAA
CAAAGGGTTAGGGCTAGGAGAGGTCAAGCTACTGACCCTCACAGCATAGCTGAAAGATTGCGCAGA
GAGAGAATTGCAGAGAGAATGAAAGCTCTTCAAGAACTTGTACCTAATGCCAACAAGACAGATAAG
GCTTCGATGCTGGATGAGATCATTGACTATGTCAAATTCCTCCAGCTGCAAGTGAAGGTCCTGAGC
ATGAGTAGATTGGGCGGTGCTGCTGCTGTCGCACCCCTTGTTGCTGATATGTCCTCAGAGGGAGGA
GGAGGTGATTGTATCCAAGCAAACGGGCGGAACCCTAACGGAGCTCAGACGACGTCAGCTAACGAC
AGCTTAACTGTGACGGAACACCAAGTGGCTAAGCTGATGGAAGAAGATATGGGGTCAGCCATGCAG
TACTTACAAGGCAAAGGGCTTTGTCTAATGCCCATCTCACTTGCAACTGCTATATCGACTGCCACG
TGTCACTCTAGGAACCCCATCATCAGCACCAGCAACAACAACAATAACAACGGCAATCCCCACCAC
AATCCTTTGCTTCAGTCCAATGGTGAGGGCCCGACCTCACCTAGCATGTCAGTGCTGACTGTCCAG
TCAGCAACTATGGGTAACGGTGGGGCTGACGGCTCCGTCAAAGATGCTGCTTCCGTTTCCAAGCCT
TGATCAACGGCGTCGACCGACTTTATGATGCGGCTCGTCCAAAGTCTACGAAAATAAGGCGTCACT
AGATGTTGGACCTTCACGACGTCGTATTTGTGTAGACTTTGTGACTGAATCTAACAAACAACAAAA

FIGURA 24 CONT.

AAGCAAGAAAGAAGCCAATCCCAAGAACAAAAACCTTATTTTATGTTATTATTATTATTTTAATTT
GCGGGTTAGGTTTTGTTAATTTTACTATTATTAAATTTTTAATTTCTCAGATGGGATATCTCCCTC
TTTTTCTTTTATTTAATATTCTCTGTCATGATTCTCTATTAAATAATTATTTACAAAAGCCTTTAA
TC

SEQ ID NO: 340, E. grandis WO2005001050 seqID77

AGCACTTTCACTTATCAGAGGCCGAAAGCAAAACCCCCTGAACACTTCAAAGAAAAGAGGTTTTTT
CAGACACCCTCTCGCCCTCTCTCTCTCTCTCTCTCTGAGCTGAGAGACAATACCAGCTTATAAAAC
CCGAGACCCATAAAACCTCCCATCGGAATTCCCGAATCACGTTACCGGGTCATCGGGTTTCGGATC
TGCCACCCAGTTCACCAACACCACCACCACCACCTTCCTTCCCCCCCTCCTTTTTGGGTTTCCCAA
TGAGCACATCTCACCTCCACTTCCTCCCTCCCCTTCCCTTAAACATTCTCTCCCACCAATAATTTC
CACATTGGCAACCAATTTCTTATTCATCATCCCCTCCCCGAGTCCCACCAACCTTAGGCTCTCTCT
CTCTCTCTCTCTCTGCCACTCAATCTTTCCACTTAGAAAGAAACCAACTTTTTGCTGGGTTTCGCA
TAGAGCAGAAAACCCATCGTAGACTTTGATCTCCAGTTGACACTTCTATAGAGAAAGGGAGATAGA
GTGGAGAGAGAGAGAGAGAGTTTATAATCATCGGCTTATGGCAGGCTCAAATCCGCAGGAAGGACT
GGGGGATGAGTTCTTTGAGCAGATCTTGGCCGTGCCGCCCGGAGCTTACTCCGGTGGCTCCACGCC
CATGGTCTTGCAACTCGGCTCCGGCCGCGGCGGCGTGGAAGATGGCGGCGGCAGAGGAGGCGGCGG
GGGGCTGAGGGGCATGGGGGTGATGATGCCCCTGGGGCTGAATTTGGAGCAAGGTGGATTGTTTAG
GCACGAGGATGTTGAGAACAGTACTAGTGCATCTTCCACCACTTCTGCCATCAATATGGAGAGAGA
AGCAGGAGTCCACCACGGCAACAACATGACAAGCTGCTTGTTTCCAGCATTTGGACAGTTCCAGAC
TCATCAGTCAGTCCAGCCACCTCCTCCCCATCCACCCCCTCCTCAACTTCACCCGGCCTTCTTGAA
TCAGCCCGCAGTGGGATCGCCGAACCAGCCGGCGATACGCCCAAGGGCACGAGCTAGACGAGGGCA
AGCCACGGATCCTCACAGCATCGCTGAGCGGTTGCGCCGAGAAAGAATTAATGAAAGAATGAAGGC
GTTGCAGGAGTTGGTTCCCAGTTGCAACAAGACGGATAGAGCAGCCATGCTTGATGAAATCGTGGA
TTACGTGAAGTTCTTGAGACTCCAAGTCAAGGTCCTAAGCATGAGTAGATTGGGCGGAGCTGGTGC
TGTGGCTCAACTTGTAGCTGATGTACCCCTCTCATCAGCTGAGGGCGAGATCATCGAAGGCGGGAA
CAATCAACCGGCGTGGGAGAAGTGGTTGACAGACGGCACAGAGCAGCAAGTGGCTAAGCTGATGGA
AGAAGATGTTGGGGCTGCAATGCAGTTCCTCCAGTCCAAGACGCTCTGCATCATGCCCATTTCGCT
CGCCTCCGCAATCTTCCGCACAAGCCAACCAGACATGCCCCGGTCCATCAAGCCCGAATCCAGCGC
CCCTTCCTGAAGCACCCTGAACAAGTTCCTGGCGACATCATAGTTGCAACAATGCCTAGGTTTCTC
AGGTCTGACCGGTTCGAGGTTGTCCCCAAGTCGACGTAACTCCTTTGCCAACCAAGTTGCTTAGTG
CCCTCTGTTCATAGTTCACTGTCGCTGTCTTAGAGTGGGTGTGAACTTGTGGTTCTCATGCCTTAA
TTACAACCTAAAAAAATGCATCAGTTTCATGTTTTTTCCAATTCATCCAAGGAAAAGTCAAAAGCA
ACTTTATTCCAAATTCAGCTTATTAAAACCTATCTTCTGATGAAAAAAAAAAA

SEQ ID NO: 341, E. grandis WO2005001050 seqID1671

CCTCCTCGCGTCCCTCTGCGTTCCGCCAGCTAACCGAAATTCCGGCCACCCCCGCCGGAGCTCCGC
CGTGAAACCCTCTCTCCGTCTCCGCCTCCGCCTCCGCCTCCGTCTCCGTCGCTTCCTTCCGCCGTC
GCGACCGTAGATAGATGCCCTTCCCGCCCGCTCATGGCGAACAACCCGAGTGACGGCGCCAGCGAC
GAGTTCCTGGAGCACCTCCTCGGCCTCCCCAACTTCGCTGCCGCGGCGGAGGCCGGCCTCGCCCCG
GGCGACGGGACCGGCCTGTCCGTCACCGCGGCCACCCCGATGATGCTCCAGCTCGGCTCCGGCGAC
GGCTCCGGCGGCCACGGCCACGGCCACGGCCACGGCCATGGCCAAGGCCACTTGTCGGCCCTCGGC
GGCGGGGGAGCGGGGGCGGGGGGCGGCGGTGGCGGCGGTGGCGGGGGGGTACCATGGGGCGGTGTTC
CCGCTGGGGCTGAGCTTGGAGCAGGGGAAGGGAGGGTTCCTGAAGCCCGAGGAGGCGTCCGGGAGC
GGCAAGCGGTACCCCGAGGAGGTCGTCGACGGCAGAGCTTCGACTGTGAAAAACCTGCACTCCCCG
TTGCCCAACTTCCCAGATGTTTCGCAAATGGCTTCTTTTAGGAACCGAGATGTGTTTCATGGACAA
CCAGTGCCCAACCCTGTTCCTGCAGCACCACATCCACCAGCAATGCGCCCGAGAGTAAGGGCTAGA
CGAGGACAAGCTACAGATCCCCACAGCATCGCTGAACGATTGCGGAGAGAGAGAATAGCAGAAAGG

FIGURA 24 CONT.

ATTCGACAATTGCAAGATCTGGTTCCAAGTGTTAACAAGACAGATAGAGCTGCCATGCTTGATGAA
ATTGTGGACTATGTGAAGTTCCTGAGGCTCCAAGTTAAGGTTTTAAGCATGAGCAGATTGGGCGCA
GCTGGTGCAGTGGCCCCACTTGTGACCGACATCCCACTATCATCAGTCGAGGAAGAAGGCGGTGAA
GGTGGAAGAAACCAGCCAGCGTGGGAAAAGTGGTCGAATGACGGCACGGAGCGGCAAGTTGCTAAG
CTCATGGAGGAGAATGTTGGCGCCGCAATGCAGTTCCTTCAATCCAAGGCTCTCTGCATCATGCCA
ATCTCTCTAGCCACCGCGATCTACCAAACTCAACCGCCCGACACCTCCTCCCTTGTCAAGTCTGAA
AGCAATCCACCATCCTAGACCAGCCACAACACAGGCACCTCCCCTTTTTCCAGGTCCCACCAAAGG
CTCAAGACTTAAATGCTCCTTGGTCCCCGCTTTAGTTTGTCCCGTTTCTACACCTTCAAAAACCAG
TTGATGGCCCTTCTTCCTTCCTCAAATTGCAATCAAAGTCAGATGCCATAGCCCTTGCTTTTCTGC
CCTCTCTTTAAGCTTTTCGCCCGGACAGGTTCAATGTCAATGCCTAATAAAATACTAGCAGGCCCT
TCATGTCTTGCCAAGTTCAACTAAAGGAGAGATATTCAAACCCTCACTTGATATGCTCTCTCTCTC
TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTACCAATAGAATAT
GCTTAGTGGAGAATGATGTGAGCGACTTTATAGTGGGTGGGGGCTCTTGTTTTTCTTGGTAGTGGG
TCGTGGGACCTTGTAAAAGGGATTGGTTGTGTAAAGTACCATTTGTATAATGTAATGCGTAATGTG
TTCGATGAGGGGGAATGGATTGCTTGGTAGAAAATTGAGAAAAAAAAAA

SEQ ID NO: 342, G. hirsutum CO123623

CGATAGTGGTGGTGGTGGGTTTAGAGGAATAGGCATGATGCCTTTAGGGTTGAATTTGGAACATGG
ATTCTTAAGACATGAAGATGGAGTTGTAGTTGACAACAACAACAACAATGCTTCTTGTTCAGCTGC
TTCTGCTGTTTCTGGAATCAGTGAGAGAGATTCTATGCATATGGTAAGCTTGTTTCCACCTTTTGG
ACAAATGCAAACTCAGCAAATCCGCGCTTCACCGCCACCTCCTCAGCCTCCACCGCAGCTCCACCA
GCCATTTCACAGCCAGCCAACATCAGGTCCAGTTGCTGCTGCACCGCACCCACCTGCCGTCCGTCC
AAGGGTACGGGCTAGGCGAGGACAGGCAACGGATCCCCACAGTATTGCTGAGCGGTTGCGTAGAGA
AAGAATAGCTGAAAGAATGAAGGCTTTGCAAGAGTTGGTTCCTAGCTGCAATAAGACGGATAGGGC
CGCTATGCTTGATGAAATTGTGGATTATGTGAAGTTTCTTAGACTTCAAGTTAAGGTTCTGAGCAT
GAGCAGACTCGGTGCAGCAGGTGCGGTGGCTCAGCTTGTTGCCGATGTACCCCTACTATCCGTTGA
GGGTGATGGTGCAGGAGGTGGAACCCAGCCTCCTTGGGAGAAATGGTCAAACGATGGCACTGAACA
GCAAGTCGCTAAGCTGATGGAAGAAGACATCGGAGCTGCCATGCAGTTTCTTCAGTCCCAGGCACT
TTGCATCGTGCCGATATCACTGGCATCTGCGATCTTCCCTGCACATCAACCTGATGCACCAACAAT
CGTC

SEQ ID NO: 343, G. hirsutum DT543504

ATGATAAGCAATGGAGGTGAATCTTCTGAGTTTCGGAGATCATTTACAGGCTTGGAAACTCTTTCT
CCAATTCCACAATTATGGCATCTACAACCTTATGACTCTGTTTCTTCCCTTCCCACTTTGGTGGGA
CAAACCATAGTAGATGATGAAGGCAATATTAACAGATTTATTGAGATTGATGAAATTCTGCAACCT
GAGAACTTATCTGCTTCCATCAATTCAAAGGGACAACAGGACATGCAAAATTCCTTTTGCTCTTCT
TTTCCCGCTGACCACCCTATAACAAAGACCATGATTGGGTTGCCATCTCTGGTGCAGGGTGCATCG
CCTAACCTGAACAATGGTTCCGACCGAACTGTGAAGCCTCGAGTAAGGGCACGTCGAGGTCAGGCT
ACTGATCCACACAGCATTGCAGAGAGGCTTCGAAGAGAGAAGATTGCTGAAAGAATGAAAAACCTG
CAAGAACTTGTACCTAATTCCAATAAGACAGACAAAGCCTCTATACTTGATGAAATCATAGGGTAT
GTCAAGTTTCTTCAACTCCAAGTCAAGGTACTAAGCATGAGCAGGCTTGGGGCAGCAGCTGCAGTT
GTTCCTCTCATCACTGATGGTCGAGCTGAGGTATCTAATGGTTTATCACTTGCACCATTAGCGGGT
CAAGGGGTTGATTTTTCACCATCTCCTGATCAAGTTGTTTTCGAACAAGAAGTGGTGAAGCTCATG
GAATCCAATATGACAATGGCCATGCAATATCTACAAAGTAAAGGTTTCTGTCTTATGCCTGTTGCA
CTTGCTGCTGCCATATCCAATGTTAAGTCATCCACATCCTCATCGTCATCATCAGTTCCTGCTTCT
GATGAGAGTAAGAAACTTGGGCTTCACCAACAGTACTCTAATCAACAACACCTGCAGCAG

FIGURA 24 CONT.

SEQ ID NO: 344, G. hirsutum TC60118

AAAAACCCAACACTCTTTTTCATTCCTCTGTTTTTCTCTCTCCCACTTTCCGTTTCTCTTTCAGAA
AAAGGAAAGTTCAAGCTAAGCTAAAACTCTGAACTTTCCCCTTCTCAGTCTCCTCAGGGACAGAGA
AAGATTAAACGAAAGGAAAAGAAAAAGACTTCTTCTTTTTTTCCCTAAAAGAAAAGTTGATCTTTT
TTTGTTTGGTTCGATTCTCTTTAGCCATTTCCTTTTTGCATGTAGATGCTCTTTGGTAGCCTTAAC
TCGCCACTTCGACTCACTTCAATCCATGGCTAATAACCCCAACGAGTCCCCCGCTGATGATTTCCT
CGAGCAAATCCTCGGACTCTCCCACTTTGCGCCTACTGAGACTGGCTTAGCGGGGCCCGACGGTAG
ATTGTCTGGAAACGCCACGACGGCTGGAGCGCCAATGCTTCTGCAGCTCAGCTCCGGCGGTGGCAC
CGGACACATCGGCGCTATAGGCGGCGGTGGAGGCGGTGCGTTTCACGGACAGGTTTTTCCATTGGG
ATTGAGCTTGGAGCAAGGAAAAGGTGGGTTTTTGAAGCCCGAGGAAGCATCCGGTAGCAGCAAGCG
CTTCCGCAACGAGGTCGTTGATGGCAGAGCTTTTTCTGTTAAAAACGTTTTCCATGGACAACCAGT
GCCAGCTACTGTTGCTGCTGGACCACATCCACCAGCAATGCATCCGAGGGTGAGGGCTAGACGAGG
ACAGGCCACAGATCCACATAGCATTGCTGAGCGGTTGCGCAGGGAGAGAATTGCCGAAAGAATCCG
GGCATTGCAGGAACTTGTTCCTAGTGTTAACAAGACTGATAGAGCAGCCATGCTTGATGAAATTGT
AGATTATGTGAAGTTCCTGAGGCTCCAAGTTAAGGTCTTGAGCATGAGTAGGTTGGGCGCAGCTGG
TGCAGTGGCACCACTCGTAACGGACCTCCCACTATCGTCAGTTGAGGATGAAAGTGGTGAGGGAGG
AAGAAGCCAACCAGCCTGGGAGAAATGGTCAAATGATGGCACAGAGAGACAAGTAGCTAAGCTGAT
GGAAGAAGATGTCGGAGCTGCCATGCAATTCCTTCAATCAAAGGCTCTTTGCGTCATGCCAATTTC
ACTGGCCACCGCAATTTACCATACGCAATCTCCGGATACCTCCTCTGTTGTTAAGCCGGAAACAAA
TCCTCCTTCATAGACACCTCACATGAGAAAAAGAGGAAAAGAAAACGGTGTATCATATGGTGGGGG
AGACTGTGGACAAAGACGGAAGCACTTTTCTTGTGTAGTGATGGGGCATGGATTTTTGGGTTTGGT
GGGATGAGGGACCCGCAAAAGGAGATGGATGAGATGAGTATATTTACATATTATAACAATAAAATA
TTACAATATTGATAGTGTTAGTGTAAGTAACAGTGTAATTTCACAGGAAAGGGCTAAAGTAGAAGA
AGGGATAATGGTTGGCTGATTTTATAAAATGCTGATTCAATAAAATAGATGGATTGGTTCACGTTT
CAAAAAAAACCT

SEQ ID NO: 345, G. hirsutum TC60119

AAAAAAAGAAGAGGAAAGCCAAAACTCCAAACTTTCTATTTTGCTTTTATATTATATTTTTTTTTG
GTTCAATTCAGTTTCGCCATTTCCTTTGCATGTAGATGCTCTTTGGTTAGTCCCTAACGCCGCTAC
TTGGACTCACTTTCCGCTGCTATCTTTTTTCCGGCTTTCTCATGGCTAATAACACCAACGAGGCCC
CCGCCGCCGATGATTTCCTCGAGCAAATCCTCGGCCTCCCTAACTTTGCGCCCTCCGAAACGGGCT
TAGCTGGATCCGACGCTGGATTGGCTGCAACCGCTGCTGGAGCTGGAGCTCCAATGTTTTTGCAGC
TCAGCTCCGGTGATGGCGCCGCTCATATCGGTGGAATCGGCGGCGGTGGAGGCGGTGCGTTTCATG
GGCAGGTTTTCCCCTTGGGCTTGAGCTTGGAGCAAGGGAAAGGCGGGTTTTTGAAACCCGAGGAAG
CTTCCGGTAGTGGGAAGCGGTTTCGGAACGGGGTCGTTGATGACCGAGCTTCTTCTGTAAAAAATG
TTTTCCATGGCCAACCCATGCAAGCAACTGTTTCTGCAGCACCACATCCACCGACAATGCGTCCAA
GGGTGCGCGCTAGACGAGGACAGGCCACGGATCCTCACAGCATTGCTGAACGGTTGCGCAGGGAGA
GAATTACTGAAAGAATCAGGGCATTGCAGGAACTTGTTCCTAGTGTTAACAAGACTGATAGGGCAG
TCATGCTTGATGAAATTGTAGATTATGTCAAGTTCCTGAGGCTTCAAGTTAAGGTTTTGAGCATGA
GTAGGTTGGGCGGAGCTGGTGCAGTGGCACCGCTTGTAACAGACATTCCACTATCATCAGTAGAGT
ATGAAAGCGGTGAGGGTGGAAGAAGCCAGCCAGCATGGGAGAAGTGGTCAAATGATGGCACAGAGC
AACAGGTAGCTAAGCTGATGGAAGAAAACGTTGGAGCTGCCATGCAATTCCTCCAATCAAAGTCTC
TTTGCATTATGCCTATCTCACTAGCCACTGCAATTTACCACACACAGGTACCAGATACCTCCTCTG
TTGTTAAGCCAGAAACAAATCCTCCTGCATAGACCTGAGGGGAAAAAAAGGGTAGGCATCCCACAG
TCCATCCTTCATCCCGCCTTTTATTTGTCCTATTCTAACACACCCACTCCCTCCCTCTCTGGCTTG
AAAATGCCATCAAAGTCAGATGCCATAGCTCTTGCTTTTTGCCTTCCTTTCCAGCTCTTTTTCGAT
TTTTCAAAACAGGGTTTGATGTCAATGCCTTAAAAAAATCAGAGACAGGCAATTGATGTCTCGACA

FIGURA 24 CONT.

AGTTCGAGTAAAGGGGAAGAATTGTCAAATTTCACTTGATATGCTCTTCTTTTTTGTATCAAAATG
GGATAGTGGACAATGATGGGAGCACTTTTGTAGTGGGGCTCTGATTTTGTGGTTAGTGGATTGTGG
GACCTGCAAAAAAAAAAAAAAAGGATAGAGATGATCCAGGGTATAATATTGTTATTGTAACAATGTA
TTTTGAAGATAAAATCGAATGCAAAAGTAGAAGAATTGATATT

SEQ ID NO: 346, G. hirsutum TC61833

AAAAAATCCCTTTGAAAAAAAAAATTCCCTCTTTTCCTTTCTGTTACTTTCACAGAGCTGAAAGTA
CAGAACAGAAAAGAAACCCCAAACCTCAAAATTTTCTCTTTCTCACGTTGATAGAAGGATTTAAGC
TAGAACGGCAAAGATATGAAATAATATCGTGTTTTTTGTCTATAAATATTTTGTTGGGTTTTCAAT
TCTTTTTTCGCCATTTCTTTTGCATGTTGATGATCTTTGGTGGCTTTAACTCGGTACTTGGACTCA
CTTCCGCTACTACTTCTTCGGGTTTCACATGGCTAATAACCCCAACGAGGCTTCCGCTGATGATTT
CCTCGAGCAAATCCTCGGATTCCCTAACTTCGCGCCTTCTGAGACGGGTTTGGCGGGACCCGACGG
TGGACTGACAGGAACAACTGCTGGTGCCGGAACGCCAATGTTTCTACAGCTCAGCTCTGGTGATGG
AGCCAGTCATCTCGGCGGAATCGGCGGCGGTGGAGGCGGCGCGTTTCCCGGACAGGTTTTTCCGTT
GGGGTTGAGCTTAGACCAAGGGAAAAGCGGCGGGTTCTTGAAGCCGGAGGAAGGTTCTGGTGGTAG
CAGCAGGCGGTTTCGCGATGAAGTCGTTGATGGCAGGGCTTCCTCAGTTAAAAACGTTTTCCATGG
TTCACCAATGCCGGCTACCGTTGCTCCATCACCGCATCCACCAACAATGCGTCCAAGGGTGCGGGC
TAGACGAGGACAGGCCACGGACCCGCATAGCATTGCTGAACGGTTGCGGAGGGAAAGAATTGCCGA
AAGAATCCGAGCATTACAGGAACTTGTTCCTAGTGTTAACAAGACTGATAGAGCGGTCATGCTTGA
TGAAATTGTAGATTATGTCAAGTTCCTGAGGCTCCAAGTTAAGGTTTTGAGTATGAGTAGGTTAGG
CGGGGCGGGGGCAGTGGCACCGCTTGTTACAGACATCCCATTATCATCGGTCGAGGACGAAAGCGG
CGACGGTGGAAGGAACCAACCGGCATGGGAGAAATGGTCAAACGACGGCACCGAGCGACAAGTAGC
CAAGCTCATGGAAGAAAACGTAGGAGCTGCCATGCAATTCCTTAAATCAAAAGCTCTTTGCATTAT
GCCAATCTCACTAGCCACCGCCATCTACCACTCGCAACCACTGGATACCTCTTCTATCGTCAAGCC
CGAAACAGATCCCCCACCATAACTCTAACAAACAAAAAATAAGGGTATGCATCCCACACAGCCCAT
CCTATATGTCCAACACTATGAGGAAAAAAAAAGGAAGGGGGGGCCTTCATGTCCGGATTTGTCAAT
GGTGATCTCTTTTGTAGTCAATGGTGGGGGTAGGGGATAATGATGGAAGAAGAGGGGCACTCTCAT
AGTGGGGGCTTTAAATTTTTTTGTT

SEQ ID NO: 347, G. hirsutum TC67603

CAAAAAGAAAAAAAGAGAAACTCGTCCCTTCATCTTCAATCTCATTCATCATCAAAGCAAAAAAAA
AACCCTAATTATGCAGCCTTGTAGTCGTGAAATGCAAGCAATGAACTCTCTCTTAAACCCGACCCA
ACAAATCCCTCTCCAAGACCTTCAAATCAACGGTAACAGACACCACCATCAACAGATCCACGTCCC
ATCATCATCGCAGTTCCAGTATCCCACACCCACTTCAACAACCCACCATGACGACAGCTTCTTAGA
ACAGATACTCTCCTCCACCACGTCCTTCCCGTGGTCCGACGAAACGGGTCCTCCCAACCCCGAGGA
TAACAATAACGCCGGCTTCAACTACGATGAGATGGTTCTAGCTTCCAAGCTTCGACAGCATCAAAT
CAACGGCGGAGCTGCTGGTAACCCTTTTGCCGCTATGAAGATGATGTTGATGATGCAACAACAACA
ACAGCAACAGCAGATGATGTTAGCGGGAAGACCCACGGTTGCCTCCGCCGCCGCCGGAGGCGGAAT
CACCACCAATCATCAGGGTGGAGAGGGTTCCATGCAAGCTTTGTATAATGTTTTTGGCGATGGATC
TTTGCATGGAACTATGCAGGCGAAAAGCTTTGTGGGGGCTAATTCAGCAACAGAGATGACCCAAAG
TCAGGGAAGTGGTCCGACGGCCGGAGAAGCGGTGACGGCGCCGGAAAAGCCTAAACAAAGAGTTAG
GGCAAGGAGGGGTCAAGCTACTGACCCCCACAGTATCGCGGAAAGACTACGTAGAGTTAGAATTGC
AGAGAGAATGAAAGCTCTTCAAGAACTGGTTCCCAACGCCAACAAGACAGATAAAGCTTCAATGCT
TGATGAGATCATCGACTATGTCAAATTCCTCCAGCTCCAAGTGAAGGTTCTGAGTATGAGCAGATT
GGGTGGTGCTGCTGCTGTTGCTCCCCTTGTTTCTGAGGGAGGTGGCGATTGTATTCAAACAAGTGC
CAATGGTGGGTCCCATCCACGAAATTCCAACGGCGACCAAACGCCGTCGGCCAACGACAGACTGAC
GATGACGGAGCACCAAGTGGCCAAGCTTATGGAAGAGGACATGGGCTCCGCCATGCAGTATCTGCA
AGAGAAGGGTCTGTGCCTCATGCCGATCTCTCTCGCCACCGCCATCTCAAGCGCCACGTCTCGTTC

FIGURA 24 CONT.

AAGGAACCCCATGATCAACCATGAAAACCCGGCCAACGGCAGCCACCTTTTGATTCAATCAAGCGG
CGGCGATGGACCTTCCTCGCCTAGCATGTCCGTGTTGACAGTCCAGTCAGCCACCATGGGTAACGG
TGGCCTTGAAGGTGGCGCAGCCTCGGTTTCCAAGCCCTGATAACGGCGGTGACTGACATACGTAGA
GAGGAATTGAAGGGAAACAGCTTTTAGTCTATGAGAATAAGGCGGTCTTCTTCCCAAGTATTTTCT
TAAAAGTACGAAACGACGCCGTATTTGACGTAGACTTATAAGCTAACCCCTAAGACTGACCTCATC
CAATCAAC

SEQ ID NO: 348, G. max TC229602

GGATTCATGAAGCCCGACGAAGCCTCCGCTAGCGGAAAGCGCTTTCGCGACGACGTCGTTGATAAT
AGAGCTAAGCATGTTTTTCATGGGCAACCCATGCCTACTACTATGCCTGCTGCTCCTCATCCTCCA
GCAATACGTCCTAGGGTGCGGGCCATAAGAGGACAGGCTACAGATCCACACAGCATAGCTGAACGA
TTGCGCAGAGAAAGAATAGCAGAAAGAATCAGGGCATTGCAAGAACTGGTTCCAAGTGTCAACAAG
ACAGATAGAGCTGCCATGTTAGATGAAATTGTGGATTATGTCAAGTTCTTAAGGCTTCAAGTGAAG
GTTTTGAGCATGAGTAGACTGGGTGGAGCCGGTGCAGTGGCACCACTGGTAACTGACATCCCATTA
TCATCAGTCGAGGAAGAAGGTGGTGAAGGTGCGCGAAACCGGCCAGCATGGGACAAGTGGTCAAAT
GATGGCACAGAAAGACAGGTAGCTAAGCTTATGGAAGAAAACGTTGGGGCTGCCATGCAGTTTCTT
CAATCAAAGGCTCTCTGCATCATGCCAATCTCACTGGCATCGGCAATATACCAGTCACAACCACCG
GACACTTCTAGCATAGTCAAGCCTGAAACTAATCCTCCTTCATAGACCTAATTCACATGTTGCACA
ATCTATTATTATTGGTCCCATCCCAGGCTCACCTAATACTTGCTGGGACCCACACCTAACGCGTTA
TTTGTCTGTTCCAAATCCCCCCCCCATAAAATAAGTCATTGGCAAGCATCAAGTTGCAATCAAAGTC
AGATGCCATAATTCTTCCTTTTTTTGCCTTTCTTTCAGCTTTTGTCAATACAGGGTTCGATATCAA
TGCCTTGGAAAAGACAAGCAAACCCCTATATCATGTCTTGCCGTATTTCAAGTAGAGAGGAAGACA
TTGTCAAAAGTTCAATTGATATGCGCTTCTTTTCTAGTGTCAAATGGGTGGGGGTAGTGGACCATG
ATGTAGGCGCTTTTATAGTGGGGCTTAGTTTTTGTAGTGGGCTGTGGGACCTTCTAAAAGGT

SEQ ID NO: 349, G. max TC205173

GTCCAACAACAACAACGACGCTACTACTAATGTCGCATCTTTTTCCTCCTTCGATGAGCACTCCAC
CTTAGCCTCCAAGTTCCGCAACCACCAGATCAGCTCCAATAACAACGCGCCCAAAAACGCCGCGGC
GGCGGCTCTCATGCTCCAACACCAACTCCTCAGAGACTCTGGTCTTCTCAACATGCCTCTTTCCTT
GCCCGGAAACGACGTCGTGGTTGACGCTTCTACCTTCGAATCCCCCAATCCGAGTGGTAAAGCTTC
AGTCCAAACTCTCTACAACGGGTTCGCCGGATCTCTGCATGGCGTAGGCCAATCCTCGAACCAGAC
TCAACATTTTCAACACCCTCAGGGAAGTTCGAATCCGATGCAAGGGCAAAATTTTGGGGCGGTGCC
AGCTGGTGGTGGTAGCGCGACGAATCAGGCTCCAGCGAGTGGCGCTGCCGCGGGTGGCGCGCCGGC
TCAGCCGAGGCAGAGGGTTCGAGCGAGGAGAGGTCAAGCCACGGACCCACACAGCATCGCTGAAAG
GTTACGTAGGGAGAGAATCGCCGAGAGAATGAAGGCCCTACAAGAACTGGTTCCCAATGCCAACAA
GACAGACAAGGCATCCATGCTAGATGAGATCATCGATTACGTGAAGTTCCTGCAGCTCCAAGTCAA
GGTTCTGAGCATGAGCAGATTGGGCGGTGCTGCTGCTGTCGCTCCCCTTGTTGCTGATATGTCCTC
TGAGGGAGGCGGAGACTGCATTCAAGCCAACGGCAAGAGTAACGGCGGTGGGGCCCAGGCTTCAAC
CACCAACACCAACACCAACCAAACGACAGCGACAACAACCTCAAACGATAGCCTAACGATGACGGA
GCACCAGGTCGCGAAACTAATGGAAGAAGACATGGGTTCGGCTATGCAGTATCTTCAAGGAAAAGG
TCTTTGCCTCATGCCAATTTCACTTGCCACTGCAATCTCCACTGCCACGTGTCCCACCAGGAACGT
TAACGTTAACCCCTTGATCAACGCCGCCGCCGGAGCCACCCATTTTCCGACCGCAGCCAACCCCGC
CGGCGAAGGGCCGTCGTCTCCGAGCATGTCCGTACTGACGGTGCAGTCCGCCATCGCAGGCAACGA
CGGCGCCGCCTCCGTTTCGAAGCCGTGATGACGGTGTTTGAGTGACTTTATAGGAAAAAGAAGAAA
AAAACAAAGTGAACAAAGGACAGAAAATGACGTGTGTTTTTTTTTCTTAAAAAAAAAGGAGCCGTT
AAAGTCTACGAAAATAAGGCGTCAGGGATACGGAGGACTCGACGACGCCGTATTTACCTAGACTTT
GTTTGGCTGATGCTAACTTTCAACCAACCAATTACAAACTCCACCATGCCAATTACCTACTTCTTC
GTCTTCTTTTTCTTTATTATTATTGTTAATTGTTAACTGTCGCTTTGGGTTTCTTCTGTTCATTGT
TTTTAAAAATCTGGAGTGGAATTTCTCTTTTTGGTGTT

FIGURA 24 CONT.

SEQ ID NO: 350, G. max TA55042

TTTTCTTCTTCCTTTTCGTGCTCTCTCTCAAATTCAACACGAGAGACACAAAACAAAACATTCTCG
GGAAGCGTAACCGCCCAGAATTCCGACCATTTTCACAGACCTTGTTTTCTGCTCTTGATGTTTTTC
ACCCCCCAAATGTGGCAATAAAAAACGGGTGGCTTAATTTGACACAATCGGTGGTTATTGCAAAGT
TGAGATGGCAGGTAATAGTGGTTCAGAGGGGTTGGGCGATGATTTCTTCGAACAGATCTTAGCAGT
GCCTGAAGCCGGCACAGTGGGAATGTTGCAATTGGGGTCCACAACTGGTGCTTTCAGAGGAGCTTC
TGGGTTAATGCCTCTGGGCTTGAATTTGGAACAAGCTGCCTTCCTAAGACACCAAGTTAACGTTGA
CGACGACGTTGTTCATGTTAATGTTGATGATGCCACCATCCACCAACATCACCTCACTCTCCATAA
CAACAACAACTCCTCATCACCCTCTTCCACTGCACCAATCACCGATAGGGATTCTATGCATATGAG
GGGTTTATTTTCCGCGTTTGGACAACTGCATACTCCTATCCGCCCCACGCTGCCATTGCCTCCACC
ACCTCGTCAGCCACAACTCCACCTCCACCATCATAACCAATTTCAGGGCCAGGCAGCTGCAGCTCC
AGCGTCAATGGCTGCTATGCCACAACCACCTGGGATCCGCCCTCGTGTGAGAGCAAGAAGAGGGCA
AGCTACAGATCCTCACAGTATTGCTGAGCGGTTGCGTCGTGAAAGAATCGCTGAAAGAATGAAGGC
ATTGCAGGAATTAGTTCCCAGCATCAATAAGACGGACAGAGCGGCGATGCTTGATGAAATTGTGGA
CTATGTGAAGTTCCTTAGGCTTCAGGTTAAGGTCCTGAGCATGAGTAGACTGGGTGGAGCTGGTGC
TGTTGCTCAGCTTGTGGCTGATGTTCCCCTTTCTGCAGTGGAGGGGGATCAGGACATAGAAGGTGG
AGCCAATGAGCAAGCCTGGGACAAGTGGTCCAATGATGGCACAGAACAACAGGTGGCTAAGCTTAT
GGAAGAAGATGTGGGAGCAGCTATGCAATTTCTTCAGTCCAAGGCACTTTGTATCATGCCCATATC
TCTAGCCTCAGCCATTTTTCGTATGCCTCAATCAGAAGCATCAACAGGCATCAAGCCTGAATCTAA
CAGTCACTGATATTAGCGATAGAACCGAAACAAAAAGTATCATGACAGACACTAGATAATTAAATT
TCTTCATCCTTAATTTATGACAAGGGTAATATTCTACTTTCTTGTATAATTTAATTTTCCTGATTG
GATGGAGAAGGATCTAGTCGGTGCAGCAGCTGGTTCACTCTTCTCAAGTTTTGTCAAAGAAGGAGT
AATTGTTATTCCTATTAAGTAAAATCGGAGGAGCCAAGCAAATGTCCTTTTCATGGTGAAGATTAA
TATTTTCAATGGCCTTATTCAAGTTTTTCTTGCAAATATCTGTGGTGGGGTTGGTGGTGGATGACC
AAAATATTTGTAGTGGAACCTCTTTTCTTGGTTTAGTATTTTTAGTCAGTGGGGAGTGGGACCATT
TAAAAGGAAAGGATGGTTACGAATGATGTATCGGGAGAAAACTTTAATACAAGTTGAATTATAATA
TGGTTAGTTACCTTGTAAAGATGTTTTGGAAGCATCTGACCATTGGAAAAGCTAAAGATATTTTAA
AATAATCACACATTTTGAGTCTTGTTTGCTTCTAAAACAAGTGATTTCCAGCACAAAATTAGGTGA
GAGTACTGAACCGACTGCTTGCTGAGAAAGTGGAAAATGACAGTTTGTTAATTTACTGAACAAAAA
AA

SEQ ID NO: 351, G. max TC207545

CGCGTCCGCTCCGCCGACGCCTCTCCGATGATGCTTCAGCTCAACTCCGGCGACGCCGCCACCCAC
CTAGCCTCCTTCCACGCGCCGCCCTACCAGCTCGGCCTCAGCCTGGACCAAGGTGAGGGACCCTTC
CTCACGCCTGAGGACGCTTCTGGAAGCGGCAAGCGCTTCCGCGACGACGCCGTTGATACCAGACCT
AACAACGTTTTTGATGGGCAACCCATGCCTACAACTGTTCCTACTGCTCCACATCCACCAGCAGTG
CGACCTAGAGTGCGGGCTAGAAGAGGACAGGCTACAGATCCACATAGTATAGCTGAAAGGTTGCGC
AGAGAGAGAATAGCAGAAAGAATCAGGGCTTTGCAAGAACTGGTCCCTAGTGTCAACAAGACAGAT
AGAGCTGCCATGCTGGACGAAATTGTGGATTATGTCAAGTTCTTGAGGCTTCAAGTGAAGGTTTTG
AGCATGAGTAGATTGGGTGGAGCTGGTGCAGTGGCACCACTGGTAACTGATATCCCATTATCGTCA
GTCGAGGAAGAAGGCGGTGAGGGAAGAAACCAGCCAGCCTGGGAGAAGTGGTCAAATGATGGCACA
GAAAGACAGGTAGCCAAGCTTATGGAAGAAAATGTGGGTGCTGCCATGCAGTTTCTTCAATCAAAA
GCACTCTGCATCATGCCCGTCTCACTAGCTTCAGCAATATACCAGTCACAACCCTCAGGCACTTCT
AGTATAGTTAAGCCTGAAACTAATCCTCCTTCATAATCATAGTCAGAACTCTGGCACAGATGCACC
ATCCAGTGGTCCCATCAAGGGCTCACCTAACACTGCTAGGACCCACGTGTAATTTGCCTGTTTCAA
ACCCCCTTAACCTTAGTCATTGGTTCGCATCAAGTTGCATTCAAGGTTGGATGCCATAGTCTTTCC
TTTTTGTCGTTCTTTCAACTTTTTTGTCTAGACAGGGTTTGATGTCAATGCCCTTCAAAACACAAG
CAAACCCCATCTTGATAAATTTCATGTAAAGGGAAAGAAATTGTCGAAGTTCACTTGATATGCTCT

FIGURA 24 CONT.

TCCTTTTTGTATCAATTAGTGGGGGTTAGTGGAACGTGATGGAAGCGCTTTTATAGTGGGGCTTGG
TTTTTTGTAGTGGGCAGTGGGACCTTCTAAAAGGTTAGGGTGGTGTGTATATTTTATTAGTGATGT
AACTCTTAGGGAATTTGAATTGCAGAGCTGAAATGGATAATGATTGCTTTAACAATGCTGGGTGCT
GCTAGTCCACCAGAAAATGATATTGCTTTTTTTTTTTTTATTTCAGGTGTGCTCTAGTTAGTTATT
TAAAAGTACTTTCGACCCCTTTTTTCTCATGCAGGGAGAGATGGAAGGAGAGTATGGTGCTATTTT
AAACCA

SEQ ID NO: 352, G. raimondii TA13791

TTCCTTTGTTTTTCTTTCTCTCACTTCCCGTTTCTCTTTCAGAAAAAGGAAAGTTCAAGCTAAGCT
AAAACTCTGAACTTTCCCCTTCTCAGTCTCCTCAGGGACAGAGAAAGATTAAACGAAAGGAAAAGA
AAAAGACTTCTTCTTTTTTTCCCTAAAAGAAAAGTTGATCTTTTTTTGTTTGGTTCGATTCTCTTT
AGCCATTTCCTTTTTGCATGTAGATGCTCTTTGGTAGCCTTAACTCGCCACTTCGACTCACTTCAA
TCCATGGCTAATAACCCCAACGAGTCCCCCGCTGACGATTTCCTCGAGCAAATCCTCGGACTCTCC
CACTTTGCGCCTACTGAGACTGGCTTAGCGGGGCCCGACGGTAGATTGTCTGGAAACGCCACGACG
GCTGGAGCGCCAATGCTTCTGCAGCTCAGCTCCGGCGGTGGCACCGGACACATCGGCGCTATAGGC
GGCGGTGGAGGCGGTGCGTTTCACGGACAGGTTTTTCCATTGGGATTGAGCTTGGAGCAAGGAAAA
GGTGGGTTTTTGAAGCCCGAGGAAGCATCCGGTAGCAGCAAGCGCTTCCGCAACGAGGTCGTTGAT
GGCAGAGCTTTTTCTGTTAAAAACGTTTTCCATGGACAACCAGTGCCAGCTACTGTTGCTGCTGGA
CCACATCCACCAGCAATGCGTCCGAGGGTGAGGGCTAGACGAGGACAGGCCACAGATCCACATAGC
ATTGCTGAGCGGTTGCGCAGGGAGAGAATTGCCGAAAGAATCCGGGCATTGCAGGAACTTGTTCCT
AGTGTTAACAAGACTGATAGAGCAGCAATGCTTGATGAAATTGTAGATTATGTGAAGTTCCTGAGG
CTCCAAGTTAAGGTCTTGAGCATGAGTAGGTTGGGCGCAGCTGGTGCAGTGGCACCACTCGTAACG
GACCTCCCACTATCGTCAGTTGAGGATGAAAGTGGCGAGGGAGGAAGAAGCCAACCAGCCTGGGAG
AAATGGTCAAATGATGGCACAGAGAGACAAGTAGCTAAGCTGATGGAAGAAGATGTCGGAGCTGCC
ATGCAATTCCTTCAATCAAAGGCTCTTTGCGTCATGCCAATTTCACTGGCCACTGCAATTTACCAC
ACGCAATCTCCGGATACCTCCTCTGTTGTTAAGCCGGAAACAAATCCTCCTTCATAGACACCTCAC
ATGAGAAAAAGAGGAAAAGAAAACGGTGTATCATATGGTGGGGGAGACTGTGGACAAAGACGGAAG
CACTTTTCTTGTGTAGTGATGGGGCATGGATTTTTGGGTTTGGTGGGATGTGGGACCCGCAAAAGG
AGATGGATGAGATGAGTATATTTACATATTATAACAATAAAATATTACAATATTGATAGTGTTAGT
GTAAGTAACAGTGTAATTTCACAGGGAAGGGCTAAAGTAGAAGAAGGGATAATGGTTGGCTGATTT
TATAAAATGCTGATTGAATAAAATAGATGGATTGGTTCACGTTTC

SEQ ID NO: 353, H. petiolaris TA1140

CGGGGATCTCATCTTTTCACTCTTCTTCCCTCTCTCTCTCTCTTCTTTCTCTCTCTAGGTTTTA
GGATTCGCCGGCGATACGCTTCTCCGGCAGGCATGTCGACTGACCCGCCGGAGGGCTACGGTGACG
ACTTTCTCGAACAAATTCTCGCGATTCCGTCGTACAACATCACCGGATGCTTACCTGTTAATACCG
CTGATTCTACTGGCTCTGAAACAGTTTCCGTTCATCATCAACAACAACAACCGCAACAACAACTGC
AACCGCAAAAGTTTCCGTTAGGTTTGAGTTTAGATAACGGCCGTGAAACAATCGGAGGAGCGTTTG
CAGGACAGCAACAGCAACTGCAGCAGCAGCAGCAGCGTGAGAGAGGAGGGAGCATGAATATGAGTG
GTTTGTTTCCTCAGTTTGAGAATTTGCAGTCACACTCGCTGTTACATACTGTTCCACAGGGTTTTC
AAGGGCAGCCAACTAGCAGTACAGCAGTGACTGTGCCTCATCCGCCTACGATTCGCCCTAGGGTAC
GAGCTCGACGAGGACAAGCTACAGATCCTCATAGCATAGCTGAGCGTCTACGCCGAGAAAGAATTG
CAGAAAGAATGAGGGCTTTGCAGGAGCTTGTTCCCAGCTGCAACAAGACTGATAAGGCTGCTATGC
TTGACGAGATTCTGGATTATGTGAAGTTCTTACGACTCCAGGTCAAGGTTCTTAGCATGAGTAGGC
TGGGTGGAGCTGGTGCAGTGGCACAGCTCGTATCTGATGTCCCGTTGCAATCTGTTGAGGGGGACG
GGAGTGAAAATGGATATAATAATCAACAGCAGGGCCAGGCAGCGTGGGAGAATTGGTCAAACGACG
ACACAGAACGTGAAGTAGCAAAGCTGATGGAAGAGGATGTTGGGGCGGCCATGCAATTCCTGCAGT
CAAAAGCATTATGCATCATGCCCATTTCACTTGCTTCACTAATCTACCCTAATCACCAACCTGATG

FIGURA 24 CONT.

TTAAACCCGAGCCATCTGCTCCTTCTTAGAGAGCATCTTACATGCTTCTCTCTTATAACAACTACT
ACTACTACTACCTTTTTCTGTCATTTGGTTTTGTGCCCCAACCCCAACCCGATGTCTTCTCTAATT
GATGGACATTTTGTATGAGGAAAGTAGGAAGGAAACAGGGGATGTGTATGGCCATCTCATTGTGTC
CTGGGGATTTTATTTATTATCTAGATTAAAATGTATTCTAAGAAAATCTCTTCACATGGACAACCC
TACCCAACTTAGCTGTGTTTATTATTGTTAATATCTAAATGATTTGG

SEQ ID NO: 354, H. vulgare TC140470

CGGGCAGGCGACCGATCCCCACAGCATCGCAGAGAGGCTAAGAAGAGAGAGGATAGCGGAAAGGAT
GAGGGCCCTACAGGAATTGGTCCCCAATACGAACAAGACTGATAGGGCAGTTATGCTAGATGAGAT
CCTGGATTATGTGAAATTCCTTAGGCTTCAAGTAAAGGTTTTAAGCATGAGCAGATTGGGTGGTGC
TGGTGCTGTTGCACAGCTTGTTTCTGACATTCCACTTTCAGTTAAGGGGGAAGCAAGCGATAGTGG
GGGCAAACAGCAGATATGGGAAAAGTGGTCAACGGATGGCACGGAGAAACAGGTTGCGAAGCTGAT
GGACGAGGACATCGGTGCTGCAATGCAATTTCTCCAATCCAAGGCGCTCTGCATGATGCCAGTCTC
TCTTGCCATGGCTATCTATGACACACAACATTCACAGGACGGCCAACCAGTGAAGCCTGAACCCAA
CAACACTGCCTAGTATAGTAACAGCAGCTGCAAGCCGCATCCCTGCAGCTGCTAGTAGGCATGCAT
GAACCTACCATCAAAGTGTTAAGATGAAAACTGCTGGTTTTGATCCACTAACAACAAGGAGGCAGG
AAAAGACCTAAAAGATCCTCTCCACTGGAAGTTGTATCGACGGTATATGATGTCACCTACTACCCC
TTTTCCATCTTTTGTTCAAAGGTGCTTTAATTTTCAAGGATATGGTAAGCAAT

SEQ ID NO: 355, L. perenis TA3490 part

TTTCACCCTCCTCATACGCCCACTACGCTTCTTCCTTCTAAACCAGACTACGATTCTCCGGCATGG
CGACAGACCCGCCGGAGGGCTACGGAGACGACTTCCTTGAACAAATTCTCGCGATTCCTTCTTACA
ACATCGCCGGATGCTTGCCTGGTAATACCACCGATGCTAATGCTTCTGAAACGGTGTCTGTTCATC
GTCAACAGCAACAGCAGCAACCGGTTTTTCCGTTAGGTTTGAGTCTTGATAATGGCCGTGAAACAA
TCGGAGCGTTTGCAGGGCAGCAGCAGCAGAGGGAGAGAGGAGGGAGTATGAACATGACTGGATTGT
TTCCTTCATTTGAAAATTTGCAGTCACATTCGCTGCTTCATTCTGTTCCTCAGGCTTTTCAAGGGC
AATCAACTACCAGTACAGCTGTCACTGTTCCCCATCCACCTTCTATTCGCCCTAGGGTCCGGGCCC
GCAGAGGACAAGCCACAGATCCTCATAGCATAGCTGAGCGTCTACGCCGAGAAAGAATTGCAGAAA
GAATGCGGGCTTTGCAGGAACTTGTTCCTAGCTGCAACAAGACGGATAAGGCTGCAATGCTTGATG
AAATTCTTGATTATGTTAAGTTCTTACGGCTTCAGGTCAAGGTTCTTAGCATGAGTAGGCTGGGTG
GAGCTGGTGCAGTGGCACAACTCGTATCTGACGTCCCCTTACAATCCGTGGAGGGGGACACGAGTG
AAAACGGATATAATCAACCCGCATGGGAAAACTGGTCAAACGATGACACAGAACGTGAAGTAGCAA
AGCTCATGGAAGAAAGATGTGGGGGCCGCTATGCAATTTCTCCAATCAAAAGCTCTATGCATCATG
CCCATATCACTTGCTTCACTAATTTACCCGCCCCAAACACCCGACTCCAATTCCCATGTTAAGCCC
GAAACTGTCG

SEQ ID NO: 356, N. benthamiana CK293938

TAGCCGGAGATAATCTCCGGCATGGCGACAAACCAACCGGAGGGCTATGCCGACGATTTCCTCGAG
CAAATTCTCGCGATTCCTTCCTACGCTGGCTTGCCGCCGGTGGCTGACGTGGGAACTTCATCAGAA
ACGACGTCGTTCACTTCGGCATCTGCTGCTGGTCTACAGCAACCGTTGTTCCCGCTGGGACTAAGC
TTGGAAAACGGCCGTGATGACGTCAGCGATGCTGGTGCTTATGCAGTGAAGCATGAAAGAGAAGGA
ATAAATATCGGGAATTTATATGCGGGTCTAGAACATTTGCAATCTCATGCGGTTCGTCACGCTGTG
CCTCCTGTTCACCATATCCAGTCTTTCCAAGGCCAACCAACAACAAGCACAACGATGACTGTACCG
CACTCACCAGCAATTCGTCCTAGGGTTCGAGCTCGGAGGGGACAAGCCACAGATCCACATAGCATA
GCTGAGAGGTTGAGAAGAGAGAGGATATCAGAAAGAATAAAGGCTTTGCAAGAACTTGTCCCCAGC
TGCAATAAGACAGATAGGGCTGCTATGCTTGATGAGATTCTGGACTATGTGAAGTTCCTAAGGCTT
CAAGTTAAGGTATTGAGCATGAGTAGGCTGGGAGGAGCCAGTGCAGTGGCACAACTTGTTGCTGAC
ATTCCGTTGCAATCTATGGAGGGGGACAGTGCTGAAAGTAAATCCAACCAGCATATCTGGGAAAAG

FIGURA 24 CONT.

TGGTCCAATGTTGACACAGAACAAGAGGTTGCTAAGCTCATGGAGGAAGATGTTGGTGCAGCCATG
CAATACCTTCAGTCCAAATCACTCTGCATCATGCCCATATCACTTGCTGCACTTATCTATCCTACT
CAGCAACCGGATGACATGTCCATGGTCAAGCCGGAACCGGCAGCCCCGTCATAGACCTCCAATTTA
TTGCACGTGCCTCTGAGAACTCCT

SEQ ID NO: 357, N. benthamiana TC7102

CTTCTTTATTTCTCTCTTCTCCACAAAATCATCTATTTCTAAATTGAGTCATATACTTGCACAGAA
GTTACTAAAGATATAATGCAAGCCATGAACCAGTACGATCCGACGTCGTCTCACGACGATTTCCTC
GACCAAATTCTCTCTTCCGTTCCTTCTTCTTCACTTTGTTGGCCTGACCTTTCCAAATCCTGGGAC
CCACACCACCACCACCTCTCTCCGCCGCTGCCTCCTAACCCTAGCTCCGGCGATGACCACCAGCTT
CCTCCTTCTAATCCTCTTCAGCATTTCCAATATGACGACCAGTCTTCTTCTTTCCTTGCTGCCAAG
CTCCGCCAGCACCAGATCACTGGTGGTGGTGGTGGCGGCGGCGATGCTGTAGCAGCTGCTAAAGCA
CTTTTGCTCCAGCAGCAACTTTTGCTTTCTAGAACACTAGCCGGAAACGGCCTTAGGTCTCCGACT
GGAGCCTCCGGCGATAACGGCTTCCTGAACATGCTCGGTAATGGTGACCAAAACGACAGCGTCGGA
AATCCGGCTAATGACAGTTCCGTTCAAGCTCTTTTCAACGGATTCACTGGATCTCTTGGTCAAAAC
TCAAGTCAACCTCAACATTTTCTCCATCCTCAGGGAGGAAGGATGCAAGCGCAGAGTTTCGGAGCT
CCGGCAACGGCGCCGGCGATGAATCAAACTCCGGCAGCAAGTGGTTCAGCTGGTGGAGGTACAACG
CCGGCGGCACAGCCAAAGCAACAGCGAGTGAGAGCTCGTAGAGGACAAGCAACTGATCCTCACAGT
ATTGCTGAACGATTACGTAGAGAGAGAATTGCAGAGAGAATGAAGGCTTTGCAGGAGCTGGTACCC
AATGCCAATAAGACGGACAAGGCTTCAATGCTGGATGAGATCATCGACTATGTCAAATTCCTACAG
CTCCAAGTCAAAGTTCTGAGTATGAGCAGATTGGGTGGTGCTGCAGCTGTTGCTCCCCTAGTTGCT
GATATGTCCTCTGAGGGAAGAGGAGAAGGAAATGGAGGAAGGGGAGGAAACGGAACGGCGTCGTCT
TCAAACAATGACAGTATGACGGTAACGGAGCACCAGGTGGCTAAACTAATGGAGGAGGATATGGGT
TCAGCAATGCAATATCTACAAGGGAAAGGCTTATGCCTAATGCCAATTTCCTTAGCTACAGCTATT
TCAACTGCCACGTGTCACTCCAGGAATCCCATGATCCCTAACGGCAACGGTGGCAACAACCCACTA
CTAGGCGGAGGAGGAGGCCTCGCCACCAACGGCGGTGGTGGCGAGGCTGGTGGACCATCCTCTCCT
AAGTTGTCGGTTTTGACTGTCCAGTCAGCCACGATTGGTAACGGTGGAATTGATCCCTCCGTTAAA
GACGCTACTTCTGTTTTTGAAGCTTAAGAAGTTTGTTTGGAAGTTTTTTAACGGCGTTAACTGTTT
CACTGACCCCCTTTGCCACGGAGAAAAAAGAAAATGTGAAAAATAGTGCGTGGCCGTCAAAGTTAA
GTCTACCAAAATAAGGGTCTTTTAGTCTTTATTTTTATCTTTTTTCCAGTTTCATATGAAAAAAAA
AAAAGAAGAAGAGAGAACCAAGAACGACGCCGTATTTAGGTAGACTTGGCTGAAGTAAGCTTAACT
TAAGCTAACAGCAGGAAATGCCAATGGAAATGAAGTACTTATGTTTCTTTTAATCTTTTCTCCATA
ATGATGATCCTTTGATTTCCCTGTTCTACTTAAGTTTTTCCTTAATGGAGTATTGTTTGGAAAGGA
TATATATATA

SEQ ID NO: 358, N. benthamiana TC7103

GATCTCTCATTCATTAACGGGGAAAAGAAATGCAAGCTATGAACTCACAAATGTCACTACAAGAAG
AAAATGGAGCTTCGTCGGGCCAAAATATGAGTCACTCTCATTTTGACCCGACGTCGTCTCACGATG
ACTTTCTTCAACAGATTCTCTCTTCCGTTCCTTCTTCCTCTCCCTGGCCGGACCTCTCCAGCGGCG
GTGACGGTCATTTTGACGACCAGTCAATTTTCGTAGCTTCCAAGCTCCGGCAGAATCAGATCACCG
GCGGTGGTGGCGGTGGCGCAGCTGCTAAAGCGTTGATGCTTCAACAGCAGCTTTTGCTTTCTAGAG
GATTACTCGCCGGTAATGGTGACCAAAATGATGACGTTGGAAATCCGGAAAATGAGATTTCAGTTC
AAGCTCTTTACAATGGATTCGCTGGATCTCTTGGTCAAACCTCTAATCAACCTCAGCATTTTCACC
ATGCTCAGGGAGGATCCATGCAAGCGCAGAGTTTCGGAGCACCGGCGTCGTCGACAATGAATCAAA
CGCCGGCGGCGAGTGGTGGTTCAGCTGGTGGAGGGCAGCCAAAGCAACAGAAAGTTAGGGCTCGGA
GAGGACAAGCAACTGACCCTCACAGCATTGCTGAAAGATTACGGAGAGAAAGAATTGCTGAGAGAA
TGAAGTCGTTGCAGGAGTTGGTACCCAATGCTAATAAGACAGACAAGGCTTCAATGTTAGATGAGA
TCATCGACTATGTCAGATTCCTCCAGCTTCAAGTCAAAGTTCTGAGTATGAGTAGATTGGGTGGTG

FIGURA 24 CONT.

CTGCAGCTGTTGCTCCCCTAGTTGCTGATCGGTCCTCTGAGGGGGGAGGAGGTGATTGTGCACAAG
CAAATGGAGGAGGGAGGGGCAGTAATGGAACGACGTCGTTAGCTAACAATGACAGCATGACGATGA
CGGAACACCAGGTGGCAAAGCTAATGGAGGAGGATATGGGATCAGCCATGCAATACTTACAAGGAA
AAGGCTTATGTCTTATGCCAATTTCTTTAGCCACAGCTATTTCCACCGCCACGTGTCACTCCATGA
AACTCAACAACCCACTACTACACGGCGGCGGAAGCGGAGGCTCTGCCGCCATCAACGGCGGTGGTG
GTGGACCATCCTCTCCTAGCTTGTCGGCTTCCACTGTCCAGTCAGCCACGATGGGTAACGGGGGCG
CTTGAGTTTGTTGAAATATACTTCGTAATTTACGACGTTGACTGTTTCAATTATCTTTTCTTAGAA
AAATACCGGAAAAATGTGAAAAAAGTGGTGAAAATTTCGTTATTTTAGAGCTGAAATTTCAAGGCC
TTTTCACATGGGACAAAAAAGAAAAAGAAAAGAAAAAAGGGGTAATTGAAATGTTTAGCTCTAAAT
ATTATTGGTAATTTCTGATTTTAGTTCGTGTGATATTTTATTAATTGTATTTAACATTGAATTAAT
GATAATATGGGTGTTTTTGTTTATTTACATAT

SEQ ID NO: 359, N. benthamiana TC8633

GAAATATTTAAACAATTGCTTTTATTTGGTAACCCTATTGGACAATATATTTATTAATCAAAGCTT
TTGGTGACACCATTTGCTTTTTATTTTCTTCATCGCTCTCACTGTTCTATAGAAAGTAGCAGCAAC
CAAATAATAAACACATAGAAACACTCAAAACACACACTTCCATTTCTTTTCCTATATATATACACA
AGTACACTCGTGAGTAGTACTGAGTACTAGTTTGTACATCTAATATTACACATGTTAAAAAAACAC
TTACTAGTAATTGCTTTTCTGAGAAATTAAGGTGGTCGAGAAAAAAAATCTGTTACAGCTATGGCT
AACAACCCATCAGAAGGTCCTGATGATTTCTTTGACCAAATCTTGGGATTCCCAGCTTATAACGGA
GCAGAAACCAATTTGGCGGGAAGTGATGCCGGAGCAATTCCGCCGGCGATGCTGCTGCAGCTAAAC
TCCGGCGATGCGTCGGGTCAGTTTTCCGGAATGGGTCTTGGTGTTGGACTAGGTGGTGGTGGTGGG
TTCCATCATCCATCTCAGTCCGGTTCTTTTCCATTGGGGTTGAGTTTAGAGGGAGGGAAAAGTGGA
TTCATGAAGATGGATGATGTTCCGGCCGTACCCGGAAGGAGATTTAGAGATGATGTTGTTGATAGC
AGAGCTTCTTCTTCTGTTAAACCTGGTTTTCATGGCCAACCGATGCCTTCAATGCCACATCCCCCA
GCAATACGTCCAAGGGTAAGAGCAAGGCGAGGACAAGCTACTGATCCACACAGCATTGCAGAGAGG
TTACGTAGAGAGAGGATAGCAGAAAGAATTAGAGCATTGCAAGAGTTGGTTCCCAATGTCAATAAG
ACCGATAGAGCCATAATGCTTGATGAAATTGTAGATTATGTCAAGTTCCTACGCCTGCAAGTGAAG
GTGTTGAGTATGAGTAGGTTGGGAGGAGCTGCTGCAGTGGCACCACTTGTTACAGAAATTCCAATA
TCATCAGTAGAGGAAGAGATCAGTGAAGGTGGAAATAATCAACCAGCCTGGGAAAAGTGGTCAAGT
GATGGTACAGAAAGGCAAGTGGCCAAACTTATGGAAGAAAATGTTGGTTCTGCAATGCAATTTCTT
CAGTCTAAGGCACTCTGTATTATGCCTATTTCTCTTGCATCAGCAATTTACCACTCTCAACCACCA
GATACATCAAGTCTTATTAAGCCAGAAACTAATCCTCCTTCTTAGAGATCCTTATTAGAAAGCTGG
CTCGGACATCACGTTTAATTTTTTTTAAAAAGAAAAAGATCCTTATTAAAAAGGTGACAACATGCA
AATGACCCTGAGAATTTTCTTTGGGAGTGGATGGTCCCACAGCTATTTTTACTTATTAAGTAGTAC
TATTTAGCTTTGTTGAAAGACACGGTTCAATGTTATGTATAAAGTAACCTTTTTTTTTGTGATTGA
TATGATATATAATTCAAGCAGAAGAAGAAAAATTGTCAAAGTTCTGCC

SEQ ID NO: 360, N. tabacum TC12771

AAAACAGAGTCATTCCATTCCATTTCCGTTCCTCAAAACCCTTTATTAACTTCTTCTCGCCGGGAA
AAACCCTTTTCCTCGCCGGAGATAATCTCCGGCATGGCGACAAACCAACCGGGGGGCTATGCCGAC
GATTTCCTCGAGCAAATTCTCGCGATTCCTTCCTACGCTGGCTTGCCGCCAGTAGCTGACGTCGGA
ACTTCATCAGAAACGACGTCGTTCACTTCGTCATCTGTTGCTGCTGCTGGTCTACAGCAACCGTTG
TTCCCGTTGGGACTAAGCTTGGATAACGGCCGTGATGACGTCAGCGATGCTGGTGCTTATGCTGTG
AAGCATGAAAGAGAAGGAATAAATATCGGGAATTTATATGCAGGTCTAGAACATTTGCAATCTCAT
GCAGTTCGTCACGCTGTGCCTCCTGTTCACCATATCCAGCCTTTCCAAGGCCAACCAACAACAAGC
ACAACAGTAACTGTACCGCACCCACCAGCAATTCGTCCTAGGGTTCGAGCTCGGAGGGGACAAGCC
ACAGATCCACATAGCATAGCTGAGAGGTTGAGAAGAGAGAGGATATCAGAAAGAATAAAGGCTTTG
CAAGAACTTGTCCCCAGCTGCAATAAGACCGATAGGGCCGCAATGCTTGATGAGATTCTGGACTAT

FIGURA 24 CONT.

GTGAAGTTCTTAAGGCTTCAAGTTAAGGTATTGAGCATGAGTAGGCTGGGAGGAGCCAGTGCAGTG
GCACAACTGGTTGCTGACATTCCGTTGCAATCTGTGGAGGGGGACAGTGCCGAAAGTAGATCCAAC
CAGCATATCTGGGAAAAGTGGTCCAATGTCGACACAGAACAAGAGGTCGCTAAGCTCATGGAGGAA
GATGTTGGAGCAGCCATGCAATACCTTCAGTCCAAATCACTCTGCATCATGCCCATATCACTTGCT
GCACTTATCTATCCTACTCAGCAACCGGATGACCCGTCAATGGTCAAGCCGGAACCGGCAGCCCCG
TCATAGACCTCCAATTTGTTGCACGTGCCTCTGAGAACTCCTAATAGCCTTGTGTCTTACCCTCCT
TTGTATCCTTACCTACTATCACTTGTATTTTCCTCTGCAATTGGTGGCAATGCAATCAATGTCAGA
TGCCACAATTTTTGCTTTTAGCCCAACCATTGGCCCCTTTTATTACCTTTCCGGAGTTTTCGTCAA
ACAAGGCATGATGTCATACCTTGGGCAAAGGTAAACGTCTGAAAGCACCCAATCCATTATCATCAT
TTTGGATGAGGTGAGCAGCTTTTTCAAAGTTATCTATGATTAGCTGCTCTTCTGCCATTTTGTTTG
TGGGATAGTGGGGACTCCTTTATCTTTTGCATAGTGGATTGGTGGGGCGATGTAAAAGTAATCAAC
TGGTTGTTGTAAAGTGTAATGTAATTTAAGTATATTATTGCAATGGATAGTTGATGTGGCTCTCC

SEQ ID NO: 361, O. sativa Os02g35660

GACGCGAGTCGAGAGACGGGAGAGCAGTGCAGTGCAGTGCAGTGCCGCCTCGTCTTCCTTCCCTCC
CTCCTCATACTCTCTCTCCTCCCGGTCTTCTCCTCCCCCGCCCAAAGCGGAGCGCGGCGGTTATAT
ATATATATATATATACAGCACTTCCTCCTCCTTCCTCGCCGCACGTCGCCCTCCGACTTCGCCTCG
TCAGCCGTCGTCACCGCCCCCACCCGTGGATTGCACCGGAGATCGGAATGGGCGGCTTCGCCTACC
CCTTCACGCCGTCGCCGGCGTGGTCGCGGGACGCCGTGTTCGCTGGCTCGCCGTGGGCCGCAGGCG
GCGTCTCCTCCCTCGCCGACGCTTTGGTGAGCTACGGTGCCGTCGACGACGAGGAGGCCGCGTTCC
TCGGCAAGACCGCCGCGTCGTCTCCCAGTACGGCGAGGCTGCACGAGCAGCAGCAGCTGCTGCTGG
AGGCTGAGCTGCTGCGCCACGGCGACGGCCTCGGCTTCGCCGCGATGGACGACGACGGCGGCGCCG
CGATGCTCGGCGCGCTGGAGCCGTGCGCCATGCCGCTCACCGACAGCGGCGGCCCCCCGGTCATCT
GCAGCAGCAGCTCCAACGACAGCAGCGGGAGCGAGCACTCCGCCGCCATGCCTGCAGGAGGCGGCT
TCCTCGTCGGAGAGCAGCAGCAGCACGTGCCGCCGGCGGCCTACGCCGCGGGAGGTGTACTCCCGT
CCATGGCCGCGGGAGAAGAAACGCCGCAGAGCTTCGGCTTTGGCAGCCTCTTCAATGGCGATCTCC
TGCAGGAGGCGACCGTGAGCAAGTATCATCATCATCAGCAGCAGCAGCAGCTGGGCGTTGTGCCTT
CGTCGCAGCCGCATCATCTGAATGATGACATCGATTTCAACACTGGAAAATTGATGTCTTTTGCTT
CTGGACAGCAGCACGTTACCCCCAGCATTGACAGCCTGCAGATCGACCAGAAGGAATTCAGCAGTG
GCCTGCACCACCTCAATCTCTCCTCGCTGATTTCTGGACCACTCGCATCATTCAATGCAACACAGA
GCCATAGACAGCCTGCTGAAGCTTGCGGTGGCAAAAATGGCGGCGCCGCTCCATTTGTGAACCTAT
CTGAGGTGCTACCAAAGGGCAATGGGAGTGGTAGTGCTGGGAATGGAGCACCCAAGCCTCGTGTCA
GGGCTCGCCGAGGCCAGGCCACCGATCCCCACAGCATCGCAGAGAGGCTCCGGAGGGAGAAGATCT
CTGACAGGATGAAGGATTTACAGGAGCTTGTCCCAAACTCAAACAAGACTAACAAGGCATCCATGC
TCGATGAGATCATCGACTATGTGAAATTTCTCCAATTACAAGTCAAGGTCCTTAGCATGAGCAGGC
TTGGAGCGGCCGAAGCCGTTGTCCCTCTTCTGACAGAAACACAAACCGAGAGCCCCGGGTTTCTCC
TGTCCCCAAGATCATCATCAGGAGAAAGGCAGGCAGGAGCAGGAGCAGTAACAGGAGGGCTCCCCG
GCGATCAGCCGGAGCTCCTGGACGGCGGCGCCATGTTCGAGCAGGAGGTGGTGAAGCTGATGGAGG
ACAACATGACGACGGCGATGCAGTACCTGCAGAGCAAGGGGCTCTGCCTCATGCCGGTGGCGCTGG
CGTCGGCGATATCGGCACAGAAGGGGACGTCCTCGGCCGCCGTCCGACCGGAGAAGAAGAAGAACG
GTGATGGCGATGGCGGCGGCGACGAGGAGGACGTGAAGGGAGAGTTTGACGCACCGAGGAGGCCTC
CTGTCGGCCGGCCGAAGGAGATGAGGTCCAGAGTCTAAGCTAGAACCTCTGCTTTGTTCAGGAGAG
ATGGAGCTTAATTTAGATCCTAACACTTTGCTACTTATATGTTTATTATTATGAAATTTTAATTGC
AAGTTGCAACTGGTGATCTCGAACTACATTTTGCCGGTGATTAGAAAATTGTTATATGGAGGTTAG
TTATTAAATCATATATAATTATATTTTGTATTGTTGCGGTCCTTGCGGACAATGCCATTCAGTACC
TTGATATTTG

FIGURA 24 CONT.

SEQ ID NO: 362, O. sativa Os02g55250

ATGCAACCCAACGCCCGCCAAATGCGCGGCGGCGGCGGCGGCGGCGGCGGGCAGGACGACTTC
TTCGACCAGATGCTGTCCACGCTGCCGGCGGTGTGGTCCGAGCTCGGCTCCGGCAAGCCCGCCTGG
GACCTCACGGCCGGCGCCGTAGGAGGAGGAGGAGGAGCCTCCGATGACCACTCCGCCGCGGCGTTC
GACGACTCCGCGCTCCTCGCCTCCCGCCTCCGCCAGCACCAGATCGACGGTGGAGGCGACAAGCCC
ATCATGCTCCAACTCAGCGACCTCCACCGTCATCACGGCCTCGCCGCCGGCGATGACAGCGGCGGC
GCTGCCGGGTTCCTTCCCCTGTCGCTGTTCGCTGACCGGTCGCAGGACGACATCGACGCCGCCTTC
AAGTCCCCCAATGGCGCTAGAGGTGACCATGCGCTGTACAATGGGTTCGGGGCCGCCGGAATGCAT
GGCGCCGCCGCCATGCAGCCGCCGCCGTTCGGGCAGGGAGGATCAATGCCGGCGCAGAGCTTCGGT
GGCGGAGCGGCCGCGAGCGGCGGCGGCGGCGGCGGGTCGGCGTCGGCGGCAGCGGCGGCTGGGGCA
TCGTCGGGCGGAGGGGCGGCGGCGCCGCCGCGGCAGCGGCAGCGGGCGAGGAGAGGGCAAGCCACT
GACCCACACAGCATCGCCGAACGTCTCCGGAGGGAGAGGATAGCTGAGAGGATGAAGGCGTTGCAG
GAGCTGGTCCCAAATGCCAACAAGACCGACAAGGCATCGATGCTCGACGAGATCATTGATTATGTG
AAGTTCCTCCAGCTCCAAGTCAAGGTTCTCAGCATGAGCAGATTGGGGGGAGCTTCTGCAGTCGCG
CCTCTAGTGGCTAACATGTCATCAGAGAGTAATGGGAATGGCAATGCCACCAGCAGCAGCGGGAAC
GGCGAGGCGGCCAATGGCAGCAGCAACGGCGACAACAATGGTGGCGGCACACTGCGGGTGACGGAG
CAGCAGGTGGCGAAGCTGATGGAGGAGGACATGGGCTCCGCCATGCAGTACCTGCAGGGGAAGGGG
CTATGCCTGATGCCGATCTCTCTCGCAACGGCCATCTCCTCCGCCACCTCCTCGTCGCTCCTCCCC
CGCACTGGGGGAGGCGCTGGAGGGTCCCTACATGAAGGTGGTAATGGTACATCACCACCATTGGTG
AATGGCACCGCCACCGGATGTGACGACGCCGGAGTATTCTTTTCTGTTAAATTTGTCGTGGAGCTG
TCGTTTCTCCTGCTGAATGAAGACTGTAGAGGTAAGGAGGAATCGAAATTATTGGTCCAGAAGGGA
CCCTGA

SEQ ID NO: 363, O. sativa Os03g58330

ACCTCGCTTCACACTAGCCGAGGAGACGCGAGCGCGAGCTGCTCTCCTCTCCTCCCGCCCCGCGTC
GCCGACGTCGAGGCGGCGGTCGACGTCGACGTCGCCGGGGATGGCCGGGCAGCAGCCGCAGCAGCA
GGGCCCACCGGAGGACGACTTTTTCGACCAGTTCTTCTCCCTGACCAGCTCCTTCCCTGGCGCCGC
GCCGGGCGGCCGCGCCGCCGGTGACCAGCCCTTCTCCCTCGCGCTCAGCCTCGACGCCGCCGCGGC
GGCCGAGGCGTCCGGGAGCGGGAAGAGGCTCGGGGTCGGCGATGACGCCGAGGGTGGCGGCAGCAA
GGCGGATCGGGAGACCGTGCAGCTCACCGGACTCTTCCCGCCGGTGTTCGGCGGCGGCGGCGTGCA
GCCGCCGAACCTCCGCCCCACCCCGCCTACCCAGGTGTTCCACCCGCAGCAGTCGAAGCAGGGCGG
AGCAGCCGTCGGGCCGCAGCCGCCGGCGCCGAGGCCGAAGGTGCGAGCGCGGCGTGGGCAGGCGAC
CGACCCCCACAGCATCGCGGAGAGGCTAAGAAGAGAGAGAATAGCAGAAAGGATGAGGGCCCTACA
GGAATTGGTCCCCAATACAAACAAGACAGATAGGGCAGCTATGCTAGATGAGATCCTTGATTATGT
GAAATTCCTGAGGCTGCAAGTAAAGGTTTTAAGCATGAGCAGGCTGGGTGGCGCGGGTGCTGTTGC
ACAGCTGGTTGCTGATATTCCACTTTCAGTTAAGGGGGAAGCAAGCGATAGTGGGGGCAACCAACA
GATTTGGGAAAAGTGGTCAACGGATGGCACAGAAAGACAGGTAGCGAAGCTGATGGAAGAAGACAT
CGGGGCAGCGATGCAATTTCTCCAATCCAAAGCACTCTGCATGATGCCAATCTCGCTCGCCATGGC
AATCTATGACACACAACAAACACAGGATGGACAACCAGTGAAGCACGAACCCAACACTCCTTCCTA
GTATAGTAATAGCAACTGTAAGTCTGTAACATGAATCGTTACAACCCCTAGTAGACATCAACCTAT
CAAAGTGTAAGATAAAACAGCTCGTTTTGATTCACTAACAAGGAGGCAGGAAAAGACCTAAGAGAT
CCTCTCCACTTGAAGTTGTACCAATGGTATATGATGTCACCTTTCGTTATCATCTTTTGTTCAAAG
GTGCTTTGATTGCAAGGATATGGTAAGCAATAGCCTGCTCATCCTCCTACACCTTTCAATGCCCCT
ATGAGCGCGGGTGATTGAGAGAGCTACTAAAGAGTCTGGCATGCTCTCACATTGTCTCTTTTTCGAG
TCTGTAGTGCACCAACCATACACTTTGTGGTGGGGACTGCTTTTTTGGTAGTGGGTGTTTGGACTA
ATCAAAAGGGATTGATGTATGCCTGTGTCTGCAGCCTGCACAATGGTTGACGGAATTTGAACACAA
GATGAAATGGAAGATGATTTTGATGGTGATGTTCATTTATCTCCATTTTTCCAAAGTGGAGGCGGA
GGCTTTGGTCAATGGAGGCTGAGCGTGCGCTAAGCTGATCGCTAGTGGGTCTGCTATCTACATCTA

FIGURA 24 CONT.

CATGCACTACTGTTGTATCTATTGTGACATACTCTTGATAAGAAACTGCGGTATAAAGTTTAATCT
TTCTATTGGCGATTAATCCATAGAATTACTTTGTATTTTCTGCCATGAGTTCTCAACTACCCACAT
TATTTGTGAAGAACAGCTGTTCTTGTACTGTTCACTGCCCTCTACCACAGGGTTTCCTGAGTTTGT
GGACAAATTGTGAGAATCTACAGCAAGGACTATTGAGTTTTTTTTTGTC

SEQ ID NO: 364, O. sativa Os06g08500

ATGCAGCCGAGCAGCCGCGACACGGTCGCCGGCGGCGGCGGCGAGGGGACGCAGGATGACTTCTTC
GACCAGATGCTCTCCACGCTGCCGTCGGCGTGGGCCGACCTGGGGGGCGGCGGCGGCGGCGCGGCG
GGGAAGTCGCCGTGGGAGGTCGATCCGGCAGCGGCGGCGGCGGCGTCGCAGGTGTTCGACGAGTCG
GCGCTGCTCGCGTCCCGCCTCCGGCATCACCAGATCGGTGGGGCTGGCGGCGGCGGGGGAGAGAAG
CCGGTGATGCTGCAGCTGAGCGAGCTGCACCGGCAGGCCGGCGGCGGCGAGGAGGACGGGAGCGGC
GCGTTCTCGCCGCTGCCGCTGTTCACGGACCGGACGAACGTGCCGCCGCGGGAGGAGATGGAGGGC
GGCTTCAAGTCGCCCAATGCCGCCGCCGGAGGCGAGCACGCGCTGTTCAACGGGTTCGGCGTGCAC
GGCGGCAGCGGCGGCGCCGGGCAGCCGCCGTTCGGCCAGCTTCGGAGGGAGAGGATAGCGGAGCGG
ATGAAATCGCTGCAGGAGCTAGTCCCAAACGCCAACAAGACGGACAAGGCGTCGATGCTGGACGAG
ATCATCGACTACGTCAAGTTCCTGCAGCTCCAAGTCAAGGTTCTCAGCATGAGCCGGCTCGGCGGC
GCCGCCGGCATGGCGCCGCTGGTGGCGAGCATGTCGTCGGAGGGGAACAGCAACGGGAGCAGCAAT
GGCGGCGGTGGCAAGGCGAGCAAGGGCGGCACCGGCGGCGAGGGCGGCGGCGGCGGCGGAGGAGGA
GGAGGAGGTGGCACCGGTGGTGGGATGCGGGTGACGGAGCAGCAGGTGGCGAAGATGATGGAGGAG
GACATGGGCACGGCGATGCAGTACCTGCAGGGGAAGGGGCTCTGTCTGATGCCGATCTCCCTCGCC
TCCGCCATCTCGTCGGCGACCTCATCGGCGTCGCTCCTCTCCCGCCCGTCCATCCGCCACGCCGGC
GCGCCGCCGCAGACGATGCTCGATGCCGCCGGCCCGACCTCGCCGGCGGCGATGTCTAACGGCGAT
GACCCACGGCACGCAAAGGCGGACGGCGGCGCCGGCGGGACGCAATGA

SEQ ID NO: 365, O. sativa Os06g09370

AGAAGGGGTGTACTAGTAAACAAAACAAAAAGAATAGGTTGAGACTTTCCTGCTATCGTCAGCATT
TTCAGCCCCCCTCCCCTGGGCTGGCTCTGCATCATCATTCATCATCAGCAGCAGTCAGTCCATCCA
ACGCGTCTTCAGTCCTCACCCGCAGCAAGAAAGCCACACTTTTCCTCCTCCTCCTCCTCCTCCTCC
TCCGCCTCTTCCTCTTCCTCCCCCTTCGTTCTCCACACCCACTCTCGGGCTGCGGCTTCTCCGGCT
TCCGCGCCTCTTCCTCTGTCGGGTTAACAAAGTTACTTTCAGCGTCAGAGGGTATTTCGTACAAAA
AGGAACTGCTGTTGGTACAGGATTATCAGCTGTGAACACTTTCTGTTTCTAGAGGCCATTCTCTGT
AACACTTTGGAGCATCAATTTTGTAGCGTGCTATCATGGACTACTCTGCTGGTTCCTACATGTGGC
CTGGCAATTCAGGTTCTGAGAACTATAATTTTGTTGACGGTTCGTCGGAATCATATGCAGAAGAAG
GAAGCCTACCACCTTCAGGCTATTTCATGGGAGCTGGATCAGATCGCAGTTTAAAGATCACGGAGA
ATGAAAGGAACCCCACTATGCTTGCAAATGGATGCTTGCCATACAACACCCAGGCTCATCCATTAT
CTGGCCAGATTCTACCTAAGGGTGAGCTCCCTAACAATCTTCTGGATCTTCAACAGCTACAGAACA
GCAGCAATCTGCGAAGCAATTCAATTCCCCCAGGGGTTCTTCAGTGCAATTCAACATCTGGAACAT
TTGATGCGAAATTGGATACACCTGGTCTTGCAGAACTGCCTCATGCTTTGTCCAGTTCAATTGATA
GCAATGGTAGTGACATTTCTGCTTTTCTTGCTGATGTGCATGCCGTTTCTTCAGCCCCGACGTTGT
GTTCAGCATTCCAAAACGTTTCCTCCTTCATGGAACCAGTAAATCTAGATGCTTTCGGTTTCCAAG
GGGCACAAAATGTTGCTATGTTGAACAAAACAAGTCTTCCAAATGGGAATCCCTCGCTGTTTGATA
ATGCTGCCATAGCATCACTACATGATAGCAAAGAGTTTCTCAATGGTGGTTCCATCCCTTCGTTTG
GTACTGTCCTGCAGGCACTAGGAGCGGGTGGTTTGAAGGCTGCTCAACAGGAGCAAAATATCCGGA
ATATACCTCTCCCTACATTCACATCTGGTAGTCATTTGGCAGTTACTGATGCACAAGGGCCACCAC
TTCCTTCAAAGATACCACCATTGATCCATGACCATAATAGTGAGTACCCTATTAACCATTCTTCTG
ATGTGGAACCCCAAGCAAATTCAGCTCCTGGAAATAGTGCCAATGCGAAGCCACGTACAAGGGCTC
GCCGTGGACAGGCAACTGATCCTCACAGTATTGCTGAACGGCTTCGCCGAGAGAAAATTTCAGAAA
GGATGAAAAATCTCCAAGTCCTTGTCCCAAACTCCAATAAGGCAGATAAGGCATCAATGCTTGATG

FIGURA 24 CONT.

AAATAATTGATTATGTGAAATTTCTTCAGCTTCAGGTCAAGGTATTGAGCATGAGCAGGCTAGGAG
CTCCTGGAGCAGTTCTTCCCCTTCTTAGGGAATCTCAAACCGAGTGCCATAGTAATCCATCTCTAT
CAGCTTCAACCATTTCACAAGGGCCACCAGACATGCCAGACTCAGAAGACAGCTCTGCCTTTGAGC
AAGAGGTAGTGAAGCTGATGGAAACGAGCATCATCAGCGCGATGCAGTATCTTCAGAACAAGGGTC
TCTGCCTGATGCCCATTGCTCTCGCTTCGGCCATATCCAACCAGAAAGGCATGGCTGCTGCTGCTG
CAATCCCTCCTGAAAAATGAAAAAAAAAATGAAATCGGAAAAATATGGGGGGAACAACTGCAGAAG
ACATAATAACTCAACGGCTGCAGAAATGTGCTGGACACTGCAGGTTTTGTGAATGAAGTGAAATCC
AATCCAGGGGGTCATAAAACATAGGAAGTTATCTGACTAATTCAGGCTTATGACACAATCTACTCT
ATAGTCTATAGCTATTGTGTGCCTGTTGCACATCAGTTGATGTGTATTCTTTTGTTTTTTTTGTAT
TGCATCCGGATGCTAATTAGGTAGTAGTGTTTCCTTTGATGTACATGAGTATTATCTATGCGGTTT
TATCTAATGTGCATATAAAGTAAAATCTGGTATGTTGCTGCTA

SEQ ID NO: 366, O. sativa Os09g25040

ATGGCTGCGGCGGCGGCGGCGGGCCAGATCTCCCTGGACGACCTCCGCAACGGCGGCGGGGTCGCC
GCCAATGCCGGCGGCGGCGGCGGCGTCCACGACGACTTCCTCGACCAGATGCTCAGCAGCCTGCCG
CCCTCGGCGTGGCCCGACCTCGCCGCCGGGAAGGCGGCCGAGGACGACGCCGAGGGGATGCATCAC
CACCACCACCAGCAGCAGCAGCAGTTTGGTGGGCCGTACGACGAGTCGGCGATGCTGGCGTCGCGG
CTCCGGCAGCACCAGATCAGCGGCGGCGGAGGAGGAGGCGGCGGTGGCGCGGCCGCCGTGAAGCAG
ATGGTGCTGCAGCAGCTCGCCGATCTGAGGCAGGGGCACCACATGATGCTCCAGGGCTTGGGGGGA
CGGTCGCCGGCGGGGGGCGGCGGCGGCGGCGGCGACGGGGGCCTTCTCCTCCCGCTCACCCTCGGC
AGCGGCGGATCGGGCGGCGACGTGCAGGCGTTGCTCAAGGCCGCCGCGGCCAATTCCGCCGGAGGA
GGAGACGCCGGCGGCGTCTACGGCGGCGGATTCGCCGGCTCGCTCCATCAACAGCAGCAACATTTC
CAGCCACATCCACAGACGGCGCCGACGATTCCGACGCAGAGCTTCGGCGGCGGCGGCGGCGGAGGA
GGAGGAGGAACCGCGTCAGGCGGCGGGGCGGCGCAGCCTCAGGCCGGCGCGGCTGGAGGAGGCGCG
CCGGCGCCGCCGCGGCAGCGGGTGCGGGCACGGCGAGGCCAGGCCACCGACCCCCACAGCATCGCC
GAGCGGCTTCGGAGGGAGAGGATCGCCGAGAGGATGAAGGCACTGCAAGAACTGGTTCCCAACGCC
AACAAGTTGATGCAGACGGACAAGGCTTCCATGCTGGACGAGATCATCGACTACGTCAAGTTCCTA
CAACTCCAAGTCAAGGCAAGCACGTACACTAAATTACTAATACATGTTTTGAGCATGAGCCGGTTA
GGTGGCGCCGCAGCCGTCGCGCCGCTCGTGGCCGACATGTCCTCGGAGGGGCGAGGCGGCGGCGCG
GCGAACGGCGGAGCACCGGCGGCGGCGGGGAGCGATAGCCTGACGGTGACGGAGCAGCAGGTGGCC
AAGCTGATGGAGGAGGACATGGGCACCGCCATGCAGTACCTGCAGGGGAAGGGCCTCTGCCTCATG
CCGATCTCCCTCGCGTCCGCCATCTCCTCGGCGACGTGCCATCTCCGGCCGCCGGTGGTCGCCGCC
GCGCAGCAGTTCCCGGCCGGGCTCGGCGCCGCCGCTGCCGCCGCGCACCACCACCAACTGAGCGCG
GCGGCGGCGGCGGCGGCCATGCGCGGACACCTGCCCGGCCTGAACGCCGACGGCTCCGTGCCGGCC
TCGCCGAGCATGTCGGTGCTCACGGCGCAGTCGGCCATGGCCAACGGCGGCGGGGGCGCCGCCGAC
GGCGAGGGGTCGCAGCTCAAGGACGCCGCCTCCGTCTCCAAGCCGTGA

SEQ ID NO: 367, O. sativa Os07g08440

ATGAGGGCCCTGCAGGACTTGGTCCCCAACACAAACAAGACAGACAGGGCAGCTATGCTAGACGAA
ATCCTTGATTATGTGAAGTTTCTTCGGCTTCAAGTAAAGGTTCTGAGTATGAGCAGGCTGGGTGGT
GCTGGTGCAGTTGCGCAGCTGGTTGCTGATATTCCAATCTCAGTTAAGGGGGAGGCAAGTGACAGT
GGGAGCAAACAACAGATTTGGGAAAAGTGGTCAACCGATGGCACAGAAAAGCAGGTAGCGAAGCTG
ATGGAAGAAGACATCGGAGCAGCAATGCAATTCCTCCAATCCAAAGCACTATGCATGATGCCAATC
TCTCTTGCGATGGCGATCTATGACACACAGCATTCGCAGGACGGACACTCAGTGAAGCCTGAACCC
AACACTCCTTCATAG

FIGURA 24 CONT.

SEQ ID NO: 368, P. abies DV982110

CACAGGAGGCAGCTGTAACGGCAGCGACGGAGGAATGCTTCCTCTTCCTCTCAGTCTTGGTCAGGC
TGGAAAGTCCGGCGACATGAACGAGCCCGGATCGCGGGAAGAAATTGAAGCTTCTTTCAAGTCGGT
AAATAATGCCCGAGATTCGAACCTCGGAGGGCTATTCCAGCCGTTCGCCGTATCACCGCGGGGCGT
TCGACCGACCGGTCAAAACTTCCATGCGCAGCCCGGACAAGTGCCATTGCAAGGGTATGGTGGAAT
GCCCCAACCTCAGCATCAGAATCCACCTCCGACCGGAGTTGGTGCTGCACCACCTGTTCGACCTAG
AGTAAGGGCACGCCGTGGACAGGCTACAGATCCCCACAGTATTGCAGAACGATTGCGTAGGGAGAG
GATTGCAGAGAGAATGAAGGCTCTGCAAGAACTGGTTCCTAATTCCAATAAGACGGATAAGGCTTC
TATGCTGGATGAAATAATTGATTACGTCAAGTTTCTTCAGCTGCAAGTGAAGGTTCTAAGCATGAG
TCGACTGGGAGGTGCAGGAGCAGTTGCTCCCCTTGTGGCTGATATACCTGCTGAGGGTGCCAACGC
CCCACAAGGCACAAGAACCAACGGTAGCCAGAACTCTTCTCCGGACGGTCTAGCATTAACAGAACG
CCAAGTAGCAAAGCTAATGGAAGAAGACATGGGAACGGCCATGCAGTATCTTCAGGGAAAAGGTCT
CTGTCTGATGCCTATATCTCTGGCCTCTGCCATTAATAACTCCGGTTCGAGACCCCAGGCTCCTTC
CACCCCTTCTCNNNNNNGACTATTGGCATCTAATGTTAACGACAGGCAGGTACTCGAACCCTCCTN
NNNNTCGGCACTATCTNNNNNNNNCTCNNNNNNNNATGACATACAGCCATCCATC

SEQ ID NO: 369, P. deltoides TA3263

ATTACGGCCGGGGGGTACTCCTCACTAGTCACTGGCTACCTGTGCTCGTATTAGAAAGCAAACCAA
AGCAAAGCGAAAAGCTGAAACCTTGAAAGCCGAAGTACTTAGCTATAAAGAAAAAAGAGAAAAAGAA
AGAAAGAAGAAGGTAACCCAACAGTCCTTAACTCTCTCAATTAAAACGATTTTGCATGTAAATTGA
TGCTATTTTGTATCCCCAACTCTCTATCTCAACTCACTTCGACTCCTCCGAGTCCAACTCCAAATG
GCCAATAATCCAACCGAACCTCCAGCCGACGATTTCCTCCAGGAAATCCTCGGGATGCCTAACTTT
GCTTCGGCTGAAGCCGGGTTGGTTGGAGCCGATGCTGGCCTGGCTGGCGCCGCTGCTGCTCAAGCT
TCTATGATGCTTCAGCTTAGCTCTGGTGATGGTTCCGGCCACATCTCCGACCTTGGCGGTGCTCCT
GGAGGAGGAAGCGCCGGGTTTCACGGCTTTCCCTTGGGGCTTAGCTTGGAGCAAGGGAAGGGAGGG
TTTCTGAAGCCCGAGGAGGCGTCTGGGAGTGGAAAAAGGTTCCGTGATGAGATTGTTGATGGTAGA
GCGAAAAATGTTTTTCATGGTCAACCAATGCCGACAACGGTTGCTATAGCACCGCATCCACCAGCA
ATGCGCCCAAGGGTACGGGCTAGACGAGGTCAAGCCACAGATCCACACAGTATTGCTGAACGGTTG
CGCAGAGAAAGAATAGCCGAAAGAATCAGGGCATTGCAGGAGCTGGTTCCTAGTGTCAACAAGACG
GATCGAGCCACCATGCTTGATGAAATTGTGGATTATGTGAAGTTTTTAAGGCTTCAAGTAAAGGTA
TTAAGCATGAGTAGACTGGGCGGAGCTGGTGCTGTCGCACCACTTGTAACAGACATCCCACTATCA
TCAGTTGAGGATGAAACTGGCGAAGGAGGAAGAAACCAGCCAGCTTGGGAGAAGTGGTCGAATGAT
GGCACTGAACGGCAGGTAGCTAAGCTAATGGAGGAAAATGTTGGCGCTGCCATGCAATTTCTTCAA
TCGAAGGCTCTTTGCATCATGCCCATCTCACTAGCGACGGCCATTTACCATACACAACCACCAGAT
ACCACTAGTATTGTCAAGCCAGAAACAAATCCCCAATCATAGAACCGTTGAAAGGTGGGGGAGCAT
CCCATGGTCCTGTCTAAAAGAGTAAAGTCCAAATGCTTAGGGTCCCACGTCTTATTTGTCACGTTT
CACCCCCACAAT

SEQ ID NO: 370, P. radiata WO2005001050 seqID516

GTTGCAAAGCAATACCCAGAAATTATTCACAATGTCATTGCTGCCTCCCGGGATAGGAGGCCAGGG
CTTGAACTCTCAGGAGAACGACCCGTATCCGTACGACGATTCCCAGTTCCTAGCTAACAGGCTAAG
GCAGCATCAACTGAGCGGTAATTCTTCGATGAATCAGACGCCAAACCAGGCCTCGCAAGGGATTAA
TGAAACCAATACGTCTCCGGGCCGATCCATGGTTTTGCAGCTAAGCACAGGAAGTGCAAGCGCTTC
ACAGCTGCTGGCCTCCATGGGAAATTCGCCTCGCGGCACCGCGGTAATGACGCAGTCGCCCAGTAC
AGGAGGCAGCTGTAATGGCAGCGACGGAGGAATGCTTCCTCTTCCTCTCAGTCTCGGCCAGGCGGG
TAAATCAGGCGACATCAACGAGTCCCGCTCGCGGGATGAAATTGAAGCTTCCTTCAAGTCGGCAAA
TCATGCCCGAGATTCGAACCTCGGAGGGCTATTTCCGCCGTTCGCCGCATCACCACGGGGCGTTCG
ACCGACAGGCCAAAACTTCCATACGCAGTCCGGACAAGTGCCAATGCAAGTGTATGGTGGAATGCC

FIGURA 24 CONT.

CCAACCTCAGCATCAGAATCCATCTCCGACCGGAGTTGGTGCTGCACCACCTGTTCGACCTCGAGT
AAGGGCACGCCGTGGACAGGCTACAGATCCCCACAGTATTGCAGAAAGATTGCGTCGGGAGAGGAT
TGCAGAGAGAATGAAGGCTCTGCAAGAACTGGTTCCTAATTCCAATAAGACGGATAAGGCTTCGAT
GCTGGATGAGATAATTGATTACGTCAAGTTTCTTCAGCTGCAAGTTAAGGTTCTAAGCATGAGTCG
ACTGGGAGGTGCAGGAGCAGTTGCTCCCCTTGTGGCTGATATACCTGCTGAGGGTGCCAACGCCCC
ACAAGGTACAAGAACCAACAGTAGCCAGAACTCTTCTCCAGACGGTCTAGCATTAACAGAACGCCA
AGTAGCAAAGCTAATGGAGGAAGACATGGGATCAGCCATGCAGTATCTTCAAGGCAAAGGTCTCTG
TCTGATGCCCATATCTCTGGCCTCAGCCATTAATAACTCCGGTGCGAGACCCCAGGCACCATCCAC
CCCTTCTCTTCAAGGGCTATTGCCACCTAATGCCAACGACAGGCAGATACTGGAACCCTCGAACTC
GGCACTCTCTGCCCTCACGTCCACCTCCATGACCATACAGTCTTCCATCAGCTCGCCTGGGAGCGG
AATAACGGAGCAAGGTGATAATGCACACAAGGCAGTGAACGGAACAAGGAATCCAAAGGATTCCAG
AGAAGCCAATTCCGTTACGAAGTCCAACGGGATAGGAGGCCCAGCTCTTTCCAAGAACGCCGTTAA
AGGAGAGGACGAACCTCAAAGGAGGACGGGCCAGTAAACAGCACCACCTTGCTAACGGAAACTGTA
CCATACTTCTATTAACGCCAGATTGCATGAAATTGTTCAAGCATACGAAATTTACTTTTCACATGC
TTCAACAATTTCTAAATAGTCGTAGTCGCAGTCGCAGTCGCTCTATGGGAGAGCATTCATTCAAAA
ATCCCAAAAACACACTTCAAATTTTCCTTTCTCCTGAAGTTTGGTGTTTATTTTTATTGAATGGGT
AAAAGAATCGAGCAAGTCTATCAAAATAAGGCACGCGCCCCACTTTGATGATGCCTTATTTAAATA
GACTTGCATAAATATGACATTATGTATGGTTGACCCAAGCGCTTTGTTAATCACTACGTAAGTTAT
ATCTTAAGTTGGATAATAATTATTGCTCATGTCAGCTACACTACGAAATTCTTGTTTATTTAAAAA
AAAAA

SEQ ID NO: 371, P. sitchensis TA14134

CAGGTAAAAAGAAACTTTCTTTTTGATTCGGAATAAAATTCACACAATGTGTGAAAATTCCAAAAC
TCTCACGGTAAACGTTTTAGAATGTAACATTGACTTGCATGGGAAGGATGTGAATAATTAAAAATC
CTTCCCACACTAAAACCCACTTGATCTTTATATGCTTAATACTAATCCCGGAAGCTCTTAAACCAG
AAGATTGTACAATTGAATCTTCATACTCGAATAAAAACAACTATTTCTGGTTAAACAAAAATCGAG
AAAAGACGGACGGCCATGTTACAACCACGGCAGTTTCTCTCCTCTGGTTCTTGTTAGAAAGTGGTC
AATTATTATGCTATTTGCAGAGAACTAAGATTGCCAACAAACAGCTAGTTCATTGTAGCTGGGAGA
TTTCAGTTTAAGATGACAGGCAAAATCAACATTGCAAAAAACTCCATAAGAGATGGCTCCTAACCC
ACGTCTTTTCGTATGTGAGCCACCAGTATACTTTCCATATATTCGAGAATCCTTGAAGTTAGAAGC
AACCTTTTTTGGCCTCGGATATCTTCCCTATTCCTTTTTCCTAAGTTTCGGCAAGCTAGATATTAT
CCCAAATTCTGCTCCATTTGACTCAGCGGTGGGAACTAATTCTCTTCCTCTCTTTTCAACTGTGTT
TTCCTTTAAAAGGAAATCACTGCCAAATCCAGCAGATCCCATTCCAAGAGCACCTGAGCAACTATC
GACAGTTAATGAAGCAGGCTGATTGCCTGAGCTTTGCCTGTCACTAGTACTGTCAACAGAATTTGT
TACCCCAGAACCCAACAGTGGCTTTCCAGTTGAACTCGATATAGCAGTTGCCAGTGCAATTGGCAT
CAAACAAAGCCCTTTATTTTGCAAATACTGCATGGCTGCAGTCATATCCTTCTCCATGAGCTTCAC
AACATTTTGCTCGAAGGCTAGACCATCCTGCGATAGTGAGCCGCCAGCTTGACCCAGAGCCGCTGA
TGCCAGGCTGCCAGATCCCTCAGCTGGAACATCAGCAATTAGAGGCGCAACTGCGCCTGCACCACC
AAGCCTACTCATACTCAAAACCTTTACTTGCAGTTGTAGAAATTTGACATAGTCAATGATTTCATC
CAGCATGGACGCTTTGTCTGTCTTATTAGCGTTTGGAACAAGCTCCTGCAAGGATTTCAAATTTTT
TGCAATCCTTTCCCTGCGTTCCCGCTCAGCTATACTGTGTGGATCAGTAGCTTGACTACAAACTAA
TAGCTCCTTCAAAAGATACGGTGTGTTTTCGCCAGCAATGACACTAGAAGAACTATTCTGAACATC
ATTATTTTGAGCCTGGCCTTCTAAGTTCGGACTCATATGACTTGACGGTGTTACTGACATTGTTGA
GACTCTTGAGCTCGCTTGAAGTGTTCCCATAGAATCACCAGGCAAGCAATTTGTTGGTCCTCCAGC
ACTACTATCCAAAGATGATCTGTCGGCACCATTTATGTCTTCCCATGAAGATGATAGTAATTGATC
AAAAATGTCATCCATTACAGAACTCAAGAGCTCCACTCACTTGCAAAACAATTGCTATGAAATTGC
CCTTTGTCCAGCAATAAACTAGCTGAAACAGCCTTAGCTGATAGATTTCATTTCTTGTTGCTCTCA

FIGURA 24 CONT.

AAAATCCAAAATCAGGCTTTCCCCTCAGCACCTACATTAACCGCTAGCAGCAGCCATGTCCGCCTA
TAAAATAGCAGAATTTAACCAAAGAACGATTTCTCCGTTTCCAATTTAGCCTTCTTTAATGTCGCG
GCTTGGCAAAACTGCAACCAACATGTGCCGGCATTGAACGTTAACAGGAGAGCAGAGCAGAGCAGC
CACATTAGCATAACGTTATCAGAATTTCA

SEQ ID NO: 372, P. taeda TC68930

TTTTGGGCAAAGGTCTCACTCTGAAGACAAAATGCTAGCCAGAGAAAATAGATCTGGAGATACTGA
GCATGATGATCTTCAAGGGACCCTTTTGACATCTTATACAGGACCACAAGGAGGTCAGACAGTGCT
ACCTAGAACTCCAGGGGCACAATCCCATCAACAGAATCTTAGCACACAAACTATACAATCAGGTGA
AGGAACTTCTCTACAGCAGTATGGAAACCATTCAGCTGTTTCTCAGCTGCAGTCTGGGGCTGGTGG
TGGCAATGCTGCAAATGGAGCTGCAAGGCCACGTGTCAGGGCACGTCGAGGTCAAGCTACCGATCC
ACACAGTATTGCTGAGCGGCTTCGCAGGGAGAGGATTGCAGAAAGGATGAAATCTTTACAAGAACT
TGTTCCAAACTCTAACAAGACAGACAAAGCATCCATGCTTGATGAGATCATAGAGTATGTCAAGTT
TTTGCAGCTGCAAGTGAAGGTTTTGAGCATGAGTAGGCTTGGGGGTGCTGGTGCAGTTGCACCTCT
GATTGCTGATGTTCCATCTCAGGGATCAGGTGGTGTCATGTCAACAGCCTTGGGCCAAGCAAGTGG
GCCCTTGACTCTGTCACAGGATGGTCTTGCTTTTGAACAGGAAGTTGCAAGACTCATGGAATCGAA
CATGACCTCAGCTATGCAATATTTACAAAATAAAGGACTTTGTTTGATGTCCATTGCACTTGCAAC
AGCTATATCAAGTTCAAGTGGAAAGCCTGTACTGGCTACTGTGCCAGGAACAGGTGTTGATAGTAG
TGAGAAGCAAAATTCTGACATGCAATCCATTGTCTTACCTTTTAGTAGCAGTTCAACAACCATGGG
ACTCAGAGCTACTGCATCTGGCAGTGACGCCCCTATAAATGAGAATTCAATAAATAAAGCTAGCAT
AGAAAAAGTTGTGGCAGAAAAATCAAATGGCATTTCACCTGGGCTATCATCTGATGTGCCCAAGGT
TGGATCACAGAGCAGGGAAGAACTCTTGAAGCGAGCACAATGATTTCTATCTGCAAGTGTGTCTGA
TTTCGGTATTCAGGAGTGAGCCTCAATGCTAAGGCGGTTGGTGAGATTTTCTAGTACGCAATTTTT
TAACAGATAGGTTATATACCACTGATTATGGTTCACACATCTGGGATCGTGTGAGATTTCAATAGG
AGCAAAATGTGTAATGTATAAGTGAACCCCTGTTTTAATGAGAAATATTGGTGGAACTCTGTTGGT
GCAAGGTTGCCAACCTTTTTATTTCCACTTTCCTTTTAGGCAAATTTAAGGTTACCAAGTATTTAT
TCTAACTTGAACATGACTGGTGCTGACCATACTAGGGTCCTGCTACTGAGATACCAAGATGGCTCA
GGCATCCATTTTCCTGTTTCCCTTAGTTTTGATTTGTGGATAAGTGGATATTTATCCAAACTTTAA
ACAATTGGATGATCTTATTTCAATTTGCTTCAACAAATGTTAAATGCAATGTTCTCAATACCAAGG
TAAGTATTACCACTGGGTTGTTTTCTGTTT

SEQ ID NO: 373, P. trichocarpa scaff II. 416

TCACTTTGACTCCTCCCAGACCTACTCCAAATGGCCAATAACCCAACCGAACCTCCAACCGATGAT
TTCCTGCAGGAAATTCTTGGGATGCCTAATTTTGCTTCAGCTGAAGCTGGCTTGGTTGGAGCTGAC
GCTGGCTTGGCTGGCACTGCTTCTGTTCAAGCTCCTATGATGCTTCAGCTTAGCTCCGGTGATGGT
TCCGGCCACATCTCCGCACTCGGCGGTGCTCCTGGAGGCGGAGGCGCGGGATTTCACGGGTTTCCG
CTGGGGCTTAGCTTGGAGCAGGGGAAGGGAGGGTTTCTGAAGCCCGAGGAGGCGTCCGGGAGTGGG
AATCGGTTCCGTGATGATATTGTTGATGGTAGAGTTAGAAATGTTTTTCATGGTCAACCAATGCCC
ACAACAGTCACTGCAGCTACACATCCACCAGCAATGCGCCCAAGGGTACGGGCTAGGCGAGGCCAG
GCCACAGATCCTCACAGTATTGCTGAACGGTTGCGCAGAGAAAGAATAGCCGAAAGAATCAGGGCA
TTGCAGGAGCTGGTTCCTAGCGTGAACAAGACGGATCGAGCCGCCATGCTTGATGAAATTGTGGAT
TATGTGAAGTTTTTAAGGCTTCAAGTAAAGATATTAAGCATGAGTAGACTGGGCGGAGCTGGTGCT
GTTGCGCCACTTGTAACGGACATCCCACTATCACCAGTTGAGGATGAAACTGGCGAAGGCGGAAGA
AACCAACTGGCGTGGGAGAAGTGGTCAAATGATGGCACTGAAAGACAGGTAGCTAAGCTAATGGAG
GAAAATGTTGGTGCTGCCATGCAGTTTCTCCAATCAAAGGCTCTTTGCATCATGCCCATCACACTA
GCCACTGCAATTTACCACACACAACCACCAGATACCACTACTATTGTGAAACCAGAAACTAACCCC
CCGTCATAGAACCGTGGAAAGGGGGGAGCATCCCGCGGTCCCGTTTAAAAGTCACCAAATGCTTAG
GGTCCCACTCCTTATTTGTCCCG

FIGURA 24 CONT.

SEQ ID NO: 374, P. trichocarpa scaff 70. 65

AGTAGTATTTTGTGTCATATATTTGCATGCATGGCAGGAAATCCCCCACCTGAAGGGTTAGGAGAT
GATTTCTTTGAGCAGATCTTGGCAGGGCAGCCACCTGGTTATGGTGGTGAGGCTGTGGGGTCCACC
AGCCTGCCCATGATGGGGTTGCAGTTAGGGTCTAGTGCTGCTAATGTTGGATTAATGAGGAGTAGT
GCTAATAATAATATGGGAATGATGCCTTTAGGGTTAAACTTGGAGCATCATGGGTTTTTAAGGCAA
CAACAAGATGATGGCAGTAGTTCTTTAGATACCAACAACAGCAATAATATCAATAATGCTTCTCCT
TTTTCTATCACTTCTGCTGGATTCTTGGGGAGAGATTCTGTACACATGACTAGCTTGTTCCCAACG
TTTGGACAATTGCAGATTCATTCATCAATCCGGTCTGCACCACCACCACTTGGACCACCTCAGATT
CACCAGTTTAATAGCCAGCCAACATCAGGAGCAGTTTCAGCTGTACCACAACCACCTGGCATTCGA
CCAAGGGTGCGAGCAAGACGGGGCCAAGCTACAGATCCGCACAGTATTGCTGAGAGGTTGCGAAGA
GTAAGAATCACGGAGAGGGTGAAGGCGTTGCAGGAATTGGTTCCCACTTGCAACAAGACAGATAGG
GCAGCCATGCTGGATGAAATTGTGGATTATGTGAAGTTTCTTAGACTCCAAGTCAAGGTTTTGAGT
ATGAGCAGACTCGGTGCGGCGGGCGCTGTGGCTCAGCTTGTAGCTGATGTACCACTGTCATCAGTT
CAGGGAGAAGGCATTGAAGGAGGGGCCAATCAGCAAGCTTGGGAGAATTGGTCAAACGATGGCACT
GAACAGGAAGTAGCTAAGTTGATGGAAGAAGATGTTGGAGCTGCCATGCAGCTCCTTCAGTCCAAG
GCACTCTGCATCATGCCTGTATCACTTGCTTCAGCAATCTTCCGAGCACGCCCACCGAATGCTCCA
ACACTGGTCAAGCCCGAATCGAACCCCCCCTCATAAACCTAAGCAAGAACGTGCAAGCTTTGCATA
CACATCGATCATTCCAGGGGGCATCTGTTTCCCAAATTAAAAACTGCTCT

SEQ ID NO: 375, P. trichocarpa scaff XIII. 403

GTAGTATTTTGGTGTCATATATTTGCATGCATGGCAGGAAATCCTCCACCTGAAGGGTTGGGAGAT
GATTTCCTCGAGCAGATCTTGGCAGCGCAGCCACCTGGTTACGGTGGTGGTGGGGAGGTTATGGGG
TCCACTAGCCTGCCCATGATGGGGCTGCAGTTAGGGTCTTGTGCTGGTAATGTTGGATTACTGAGG
AGTAATGCTAATAATAACATGGGAATGATGCCTTTAGGGCTGAACTTGGAGCATCATGGGTTTTTA
AGGCAACAGCAAGATGATAGTGGTAGTTCCTTAGATACCAACAACAGCAATAATATCAATAACACT
CCTCCTTCTATAAAATCTGCTCGAATCTTGGGGAGAGATTCGGTGCACATGACAAGCTTGTTTCCA
ACATTTGGACATCTGCAGATTCATTCATCAATCCGGCCTACACCACCTCCTCCTGGACCACCTCAA
ATTCACCAGATTAATAGCCATCCAAACCCAGGAGCAGTTTCAGCTGTACCACAACCGCCTGGCATA
CGGCCAAGGGTGCGAGCAAGACGGGGCCAAGCCACGGATCCACACAGTATTGCTGAGAGGCTGCGA
AGAGTAAGAATCACGGAGAGGGTGAAGGCGTTGCAAGAATTGGTTCCCACTTGCAACAAGACAGAT
AGGGCAGCCATGCTTGATGAAATTGTGGACTATGTGAAGTTTCTAAGACTCCAAATCAAGGTTCTG
AGCATGAGCAGACTAGGTGCAGCAGGCGCTGTGGCTCAGCTTGTAGCCGATGTACCATTGTCATCA
GTTCAGATTAAGGGAGAGGGCAATGAAGGAGGAGCCAATCAGCAATCTTGGGAAAATTGGTCAAAT
GATGACACTGAACAAGAAGTAGCTAAGTTAATGGAAGAAGACGTTGGAGCTGCCATGCAATTCCTT
CAGTCGAAGGCACTCTGCATCATGCCCATATCACTTGCTTCAGCAATCTTCCGAGCACGCCCACCT
AACGCATCAACACTCATCAACACCGAATCGAACACCCCTTCATAAACCTAAGCAAGAAAGTGCATA
CTTCGTGTGCTCGTCGATCATTCCAGCAGGCATCTGGGTCCCAAATCTTGTTCTTGTCC

SEQ ID NO: 376, P. trichocarpa scaf2f8. 86

CAAGACGACAAAAAAAATTCACCTGTAAGCATGATAAGTTCAAATCAAAGTGTGGAGTCCTTGGCC
ACTCAAGACACTTCATCTGTAGTTCTTGGCAGTGAGTCAGATTATGCTGTTGATAAAGTCTTAATT
TCTGAACAAGCTCGATTGCAAAATGATTGCCAGAATTGCAATGGAAACCCCTCCCCTGATGGAATG
GCACGTGGAAACTTAAAATTTGGAAACACAGGCCTGCAATGCAATGGAATTCTTCCTACCCTGAGT
TCTCTGAACTATCCAAATCAGCTTCCAATAGTTGGTGATTTGACATCGTATCTCTCTTTTTCTGAG
GCAAGCAATGCTGGATGCAATGGAAGGGAACAATCTGAGTACCTGAGATCCTTAAAAAATCTGCAA
AACTTATCATCAATCCCACAATTGTGGCCTTCACAATCTTATGAGGGTGTTTCTTCTCTCCCTCCT
TTGATGGGACAAGACAGAATAGAGGGATCTGGTCTTCGAGGAGGGAACTTGGATGATGATATGCAT
ATTATGGGGAAGGGATACATGGGCATGGATGAAATTCTCCGACTTGATAAGTTATCTGCATCACCT

FIGURA 24 CONT.

ACTACAGAGGGGAAAGAAGATTTGCAAAGTTGTCCTTTCTCATCTGGCATAGCTGAGCCCAATGTA
AACATGTCAATGAATCAATTATCATCTATGCCTCAGACTACATCAGCTGCTCCAGTAGAAGGCTGC
AATGGAACTGGAAAAACTCGTGTAAGAGCACGACGTGGCCATGCCACTGACCCACATAGTATAGCT
GAAAGGCTTCGAAGAGAGAAAATTGCTGAAAGGATGAAGAATCTGCAAGAACTTGTACCTAATTCC
AATAAGGTTGACAAGGCATCTATGCTTGATGAGATCATAGAGTATGTCAAATTTCTTCAGCTCCAA
GTCAAGGTGCTAAGCATGAGCAGGTTAGGGGCAGCCGGAGCAGTCATTCCTCTCCTCACAGATGGT
CAACCTGAGGGTCATAATAGTTTGTCGCTCTCTCCATCAGCTGGTCTAGGAATTGATATTTCTCCA
TCTGCTGATCAGATTGCCTTTGAGCAAGAAGTACTTAAGCTACTGGAATCCGATGTGACCATGGCC
ATGCAATATCTTCAAAGCAAAGGTCTCTGCCTCATGCCCATTGCTCTTGCTGCTGCCATATCCAGT
GTTAAGGCATCGTTATCAGGTACAACTTCTGAGGAGAGGAAGAACAATGGCTACACCAGTGGTCTT
GTTAGCAGCAGCAGCAGCATAACTGGCATTGACACCCATCCAATGTCCAATGATAACAATATTGCA
ACCGGTACACTTAGCAGCAAAGGCATGATTGTCAATGGCTGCAATGAGGTTGTCAAGCAAGAAGTG
CTGAAGAACACTTGATGCATTTCTAGAGCACGGAAAGTAAAGATGTAATTTCATTATTTTAGATGG
GGTTTCATTTTTCAAGCTGAGGCAGATAA

SEQ ID NO: 377, P. trichocarpa TC63334

GGGTCTCTCCGCTCTCAGTACTCCTCACTAGTCACTGGCTACCTGTGCTAGTATTAGAAAGCAAAC
CAAAGCAAAGCGAAAAGCTGAAACCTTGAAAGCCGAAGTACTTAGCTATAAAGAAAAAGAGGAAAA
AGAAAGAAAGAAGAAGGTAACCAAACAGTCCTTAACTCCCTCAATTAAAACGATTTTGCATGTAAA
TTGATGCTATTTTGTATCCCCAACTCGCTATCTCAACTCACTTCGACTCCTCCGAGTCCAACTCCA
AATGGCCAATAATCCAACCGAACCTCCAACCGACGATTTCCTCCAGGAAATCCTCGGGATGCCTAA
CTTTGCTTCGGCTGAAGCCGGTTTGGTTGGAGCCGATGCTGGCCTGGCTGGCGCCGCTGCTGCTCA
AGCTTCTATGATGCTGCAGCTTAGCTCCGGTGATGGTTCCGGCCACATCTCCGACCTTGGCGGTGC
TCCTGGAGGAGGAAGCGCCGGGTTTCACGGCTTTCCCTTGGGGCTTAGCTTGGAGCAAGGGAAGGG
AGGGTTTCTGAAGCCCGAGGAGGCGTCTGGGAGTGGAAAAAGGTTCCGTGATGAGATTGTTGATGG
CAGAGCGAAAAATGTTTTTCATGGTCAACCAATGCCCACAACCGTTGCTATAGCACCGCATCCACC
AGCAATGCGCCCAAGGGTACGGGCTAGACGAGGTCAAGCCACAGATCCACACAGTATTGCTGAACG
GTTGCGCAGAGAAAGAATAGCCGAAAGAATCAGGGCATTGCAGGAGCTGGTTCCTAGTGTCACCAA
GACGGATCGAGCCACCATGCTTGATGAAATCGTGGATTATGTGAAGTTTTTAAGGCTTCAAGTAAA
GGTATTAAGCATGAGTAGACTGGGCGGAGCTGGTGCTGTCGCACCACTTGTAACAGACATCCCACT
ATCATCAGTTGAGGATGAAACTGGCGAAGGAGGAAGAAACCAACCAGCTTGGGAGAAGTGGTCGAA
TGATGGCACTGAACGGCAGGTAGCTAAGCTAATGGAGGAAAATGTTGGCGCTGCCATGCAATTTCT
TCAATCGAAGGCTCTTTGCATCATGCCCATCTCACTAGCCACGGCCATTTACCATACACAACCACC
AGATACCACTACTATTGTCAAGCCAGAAACAAATCCCCCATCATAGAACCGTTGAAAGGGGGGGGA
GCATCCCATGGTCCTGTCTAAAAGAGTAAAGTCACCAAATGCTTAGGGTCCCACGTCTTATTTGTC
ACGTTTCAACCCCCACAATCAAATGTTCCCCTAGTTTCAAATTGCAATCAAAGTCAGATGCCATAG
TTCTTGCTTTTGCCTTTTCTTTCCAGCCTTTTTTTTTTCAAACGGGGTTCGATGTCAATGCCTAAA
AAAATGCAGGCAGGCCCTTGGTGTCCTGCCAAGTCCAAGTGAGAGGAGAATTTTCAAAGCTTACTC
GATATGCTCTTCTTTTGGAATCAGTGGTGGGGAAGTGGAAAATGATGGAAGCGCTTTTGTAATGGG
GCTCTAATTTTTCGTAGTGGGTTGTGGGACCTTGTAAAAGAGGCAAGACTGACGGGTTATATAAAT
ATATCATTATTGTAATAGTGTATTCAGCGTAAATGGAATGCAAAGGTAGAATTGAATAATGGCTGG
TTGTTTCAATATGCTGGGTGTTTCAAGTCCAACAGAAAATGATAGCAGTGCTACTTTGTTC

SEQ ID NO: 378, P. trifoliata TA5414

CTCTCTCTCTTCTTCCTGTTCTCACCTATTTAGTATTTAACATCCATCAGAAGAATGGTTTCCATT
CTCTTGAAGAAATTCTTGACTCTTCAAATTTAAATTTCCATATATCATAAAGACTTGTTCTCAAGA
AATTAAGATGCAGCCTTGCAGCAGCGGCGGTGAAATGCAAGGAATCAGCTCGCTCTTGAGCAACGG
AAGCCACGAGCAGCAGTCGCAGATCCATCAAAGCGCGACGTTTGATCCAACTTCGCAAGACGACTT

FIGURA 24 CONT.

CTTAGAGCAAATGCTATCCTCTCTCCCATCGTGTTCTTGGACTGACTTGAAGTCTCCCTGGGGGGT
TGTTGACCTTAACCCTAATAACAATAATAATAATATTAACAACAAGCAGCCCAGAGACTTATCTGA
CGAAACGGCGCCGTCCACTACTCAAGAGAATAATGTCCCTGCAGGGTTTCAGTTCGATGAGTCCAT
GATCTTAGCTTCCAAGATGAGACAGCATCAGATCAGCGGGAATACTGGGAGTGGGAATAATAATTC
TGCTGCAGCAAAGTTTATGATGCAGCAGCAGCAGCAACAACAAATCATGATGGCTGCTGCTAGAGG
TGGCATCGGGTTGCCTTTATCACTTGGAAATGGTGGCGATAACGACATCGTTGATGCCTCATCATT
CAAGTCCCAACAGGGAGGAGATGGCTCAGTGCAAGCACTCTACAATGGCTTCACTGGATCTTTGCA
TGGCTCAACCCAGCCTCAACATTTTCATCATCTTCAGGGAGGATCGATGCCCGGGCAGACCTTTGG
GGCACCAGGGCCGGTGATGAACCAGACGCAGGCTCAGGCAAGTGGGTCAACCGGTGGTGGTGGAAA
CACACCCGCTCAGCAGCCTAAACAAAGGGTTAGGGCTAGGAGAGGTCAAGCTACTGACCCTCACAG
CATAGCTGAAAGATTGCGCAGAGAGAGAATTGCAGAGAGAATGAAAGCTCTTCAAGAACTTGTTCC
CAATGCCAACAAGACAGATAAGGCTTCGATGCTGGATGAGATCATTGACTATGTCAAATTCCTCCA
GCTGCAAGTGAAGGTCCTGAGCATGAGTAGATTGGGCGGTGCTGCTGCTGTCGCACCCCTTGTTGC
TGATATGTCCTCAGAGGGGGCAGGAGGAGGTGATTGTATCCAAGCAAACGGGCGGAACCCTAACGG
AGCTCAGACGACGTCAGCTAACGACAGCTTAACTGTGACGGAACACCAAGTGGCTAAGCTGATGGA
AGAAGATATGGGGTCAGCCATGCAGTACTTACAAGGCAAAGGGCTTTGTCTAATGCCCATCTCACT
TGCAACTGCTATATCGTCTGCCACGTGTCACTCTAGGAACCCCATCATCAGCACCAGCAACAACAA
CAATAACAACGGCAATCCCCACCACAATCCTTTGTTTCAGTCCAATGGGGAAGGCCCGACCTCACC
TAGCATGTCAGTGCTGACTGTCCAGTCAGCAACTATGGGTAACGGTGGGGCTGACGGCTCCGTCAA
AGATGCTGCTTCCGTTTCCAAGCCCTGATCAACGGACGTCGACTGACTTTATGATGAGGCTCGTCC
AAAGTCTACGAAAATAAGGCGTCACTAGATGTTGGACCTTCACGACGTCGTATTTGTGTAGACTTT
GTGACTGAATCTAACAAACAACAAGAAAGCAAGAAAGAAGCCAATCCCAAGAACAAAAACCTTATT
TTATGTTGTTATTATTATTTTAATTTGCGGGTTAGGTTTTGTTAATTTTACTATTATTAAAATTTT
AATTTCTCTGATGGGATATCTCCCTCTTTTTCTTTTATTTAATATTCTCTGTCATGATTCTCTATT
AAATAATTATTTATAAAAGCCTTTAATAATTAGATTATAATT

SEQ ID NO: 379, R. communis TA1616

CCCAACGCTGCATTTTCAGGATTTCTCTCTCTACACTACAAGAACTTGAAGCCAGTACCCATAAAA
CCATATACCATCTTGGATTCCACAACCCCTTTCCTGGTTATTGGGCTGTGGATCTGCCACCCAATT
CACCAACACTTTCTTTTTCTCGGGTTTCGCATACTTCAATTCCCCTCTTCTCTCGTCTCTTCTATT
TGTCTCTCTAAAAACATACTCTCTCCCTCTCTCTCTCTCTCTCTCTCTCCTCAGCTAATTGAAA
CATCAAAAAGGAAAAAGCATTACCATATTTTTTACACTCTAAAACATCAAACTTTATCACTAATAC
TACGACTGCTACTGCTACTCTAAAACTACTTTTGCAATATTATTGCATGGCGGGAAATCCCCCACC
TGAGGGATTAGGAGATGATTTCTTTGAGCAGATCTTGGCAGTGCAGCCAGGGTATGGAGGCGGTGA
TGGTGGTGGCGGTGGAGACGTGGTGGGGTCCACCATGCCTATGATGGGTCTACAGTTGGGGTCTGC
CAGTAGTGGTGGTGGTGGATTGAGAACTAATAATATGGGCATGATGCCTTTAGGGTTAAACTTAGA
GTTCTTGAGGCAACAAGAAGATGGTAGTGGTGCTTTAGATAACAACAACCATACCAATAACAACAA
TAATAATGTTTCATCTTCTACAACTTCTGTTATCAATGATAGAGATTCTGTTCACATGGCGAGTTT
GTTCCCAACGTTTGGACAGTTGCAAACCCAGACGCTCCGGCCTCCACCACCACCTCACCTCCACCA
GCCATTTCATGGTCAGCCAACACCAGGAGTGGTTTCTGCTGCATCACAGCCACCTGCCATTCGCCC
AAGGGTGCGAGCAAGACGAGGACAAGCCACTGATCCTCACAGCATTGCTGAACGGCTGCGTAGAGA
GAGGATTGCAGAGAGAATGAAGGCTTTGCAGGAGTTGGTTCCTACAGCCAACAAGACGGACAGGGC
AGCCATGATTGATGAGATTGTGGACTATGTTAAGTTTCTAAGGCTTCAAGTGAAGGTATTGAGCAT
GAGTAGACTGGGAGCAGCAGGTGCAGTGGCTCAGCTTGTAGCTGATGTACCGCTAGCATCAGTTGA
GGGAGAGAGCATTGACGGAGCAGCAGCAAATCAGCAGACTTGGGAGAAGTGGTCAAATGATGGTAC
TGAGCAACAAGTAGCTAAGTTGATGGAAGAAGACATAGGAGCTGCTATGCAATTCCTCCAGTCCAA
GGCACTTTGCATCATGCCTATATCACTTGCTTCAGCAATCCTTCGAACACACCCCCCTGATGCACC
ATCAATTATCAAGCCTGAATCAAACACCCCATCATAAACCTAACCAAGAAATTAATCTCCTTTCAC

FIGURA 24 CONT.

GTGCAGATGAATTGCTACCAGCAGACATCTGGGCCCCAAGTGCTGGTATAGTCCCATATAACTCAT
TCTAACTCCTCATTATAAATATATATATATATATAATAAAGAAACAATATCCAGTTGGTGCAGCTG
TTCATTTTCTTTAGGAGGAGTCATTTTCATGTCAAGTAAGAAGAATCACGGGGCACTTGATGTTTA
TAATGCTATGGTCAAGGCAAATCCAAAATATTCAAAAAGCTCCATTGATGTGCTTCTATTTCTTGT
ATCTGTGGTGGGGTTAGTGGAGAATGACCAAAAAATTAATAGATAGTGCTAAGAAACTCTTCTGTT
TCATTTCTCTTTTGGATAGTGGAAGTGTAGGGCCATCAAAAAGGGGGTTGATAGGCAGTGTGCATG
ATGTATTTGAAAAATTTGTGTGGAAGTTGAATCGGATAAATGTAGGTTTGTCGTGTAAATCCATGT
TGGAAACTTCATGCTTGACAAT

SEQ ID NO: 380, R. communis TA2825

CTCCTCTCAGGCCGACTCCTGCAAATGGCAAACAATCCCACAGAACCTCCAGCCGACGATTTCCTC
CAAGAAATACTCGGCCTACCTAACTTCGCTTCGGCTGACGCAGCCGGCTTGGTCGGAGCCGACGGT
GCCTTGGCTACTCCGATGATGCTGCAGCTTAGCTCTGGAGACGGCTCCAACCACATCACTGCCCTA
GGTGGAGGGGGAGGAGGAGGAGGAGGCGCTGGATTTCACGGGTTTCCGTTGGGGCTAAGCTTGGAG
CAAGGAAAAGGAGGGTTTTTAAAGCCTGAGGAAGCTTCGGGAAGTGGAAAGAGGTTTCGTGATGAC
GTTGTCGATGGCAGAGCTAATACTGTTAAGAACGTTTTTCATGGTCAACCAATGCCCACAACTATG
GCTGCAGCCCCACATCCTCCCACAATGCGCCCCAGGGTGCGGGCTAGACGAGGACAAGCCACAGAT
CCACACAGCATTGCTGAGCGGTTGCGCAGAGAAAGAATAGCTGAAAGAATTAGGGCATTGCAGGAG
CTGGTTCCTAGTGTCAACAAGACAGATAGAGCTGCCATGCTTGATGAAATTGTGGATTATGTGAAG
TTTCTAAGGCTCCAAGTGAAGGTATTAAGCATGAGTAGATTGGGCGGAGCTGGTGCAGTCGCACCA
CTTGTAACGGACATCCCACTATCATCAGTTGAGGATGAAACTGGGGAAGGTGGGAGAAACCAACCA
GCATGGGAGAAGTGGTCGAACGATGGCACAGAACGACAAGTGGCTAAACTGATGGAAGAAAATGTG
GGTGCTGCCATGCAGTTTCTGCAATCAAAGGCTCTTTGCATCATGCCCATCTCACTGGCCACAGCA
ATTTACCATACACAAGCACCGGATACCTCAACTATTGTAAAACCAGAAACAAATCCCCCATCATAG
ACCGTACAAACGGGTAAGCATCCCATGGTCCCATCTAAAAAGGTCGCCTAAGTGCTTGGGGTCCCA
CATTTTGTTTGTCCCGTT

SEQ ID NO: 381, S. bicolor TC104646

TACTGGACTTCCCCTTCCCCAGCAGAGCAGAGCAGACCAACCGCTTGTCGCGAGCCGCCCGAAGCT
TCGGGCCCTGACGCGTGGGCCCCGACCGCGCCCGCCCCCGCCCCCGCGTCGCCGGGCATGGCGGGG
CAGCCGCCGCCGCCCCGGGGGCTCCGAGGACGACTTCCTCGAGCACTTCTTCGCCTTCCCCTCCGCG
GCCTCCGCCGCCGCCGGGGGACACGCGGGCGCCGGTGCCGGCGGGGACCACCCCTTCCCCCTCGCC
CTCAGCCTCGACGCCGCCGCCGAGCCCAAGCCGGACCGTGACCCCGTGCAGCTCGCCGGCCTCTTC
CCGCCGGTGTTCGCTGGCGCCGGCGGCGTGCAGCAACCGCACCTCCGCGGCCCACCGCCTCCGCAG
ATGTTCCAGGCGCAGCCGAAGCCGGGCGAGGGAGGCATGGCGCCGCAGCCGCCAGCCCCACGGCCC
AAGGTGCGCGCGCGGCGGGGGCAGGCCACCGATCCCCACAGTATCGCGGAGAGGCTAAGGAGAGAG
AGAATTGCAGAAAGGATGAGGGCATTGCAGGAATTGGTCCCCAACACAAACAAGACAGATAGGGCA
GCTATGCTAGATGAGATCCTTGATTACGTGAAGTTTCTTAGGCTTCAAGTAAAGGTTTTGAGTATG
AGCAGACTTGGTGGTGCTGGTGCAGTTGCACAGCTGGTTGCTGATATTCCACTCTCAGTTAAGGGT
GAGGCAAGCGACAGTGGGAGCACACAGCACATATGGGAAAAGTGGTCAACTGATGGCACAGAAAAG
CAGGTAGCGAAGCTGATGGAAGAAGACATCGGGGCAGCGATGCAGTTCCTTCAATCCAAAGCGCTA
TGCATGATGCCGATCTCGCTTGCAATGGCAATCTACGACACCCAACATTCCCAGGATGGCCACTCA
CTGATGAAGCCTGAGCCCAACACATCCTCGTAGTATAGTAACATCAACCCTAGCGTGCATTTCAAC
CCCCTAATACTATTAGGCACCGATATATCCAAAAAAAAAAAGTGTATGATATACGGAGCGTTCCAA
TGTGCTAACCGTGAGATAGGAAAGGACCTTAAATTGGTATATGATGTAGCCTCCCCTTTCCCTATT
TTATTTCAAATCAGAGCAG

FIGURA 24 CONT.

SEQ ID NO: 382, S. lycopersicum AK247217

AGCAAAGCTTTAGGTGACTCCCTTTCCTTTATTTTCTTCATGGCTCTCACTGCTCTTTAGAAAGTA
GCAAAACACACACACAAACAACCCAAAACACACACTTCCATTTCCATTTTTTGCCTATATATATAT
ATATATATGTAACAAGTACTTGTTAGAGTACTAATTTGTACACCCTCTTATCAATTCCGTTGTAAA
GAACACTTTTTTAACTAGTATTTGTTCAATTGAATGCTTTTCTGAGGAATTAAGGTGGTCGAAAAC
AGCTCCGATCCAAGAAAGATTGAATCTTTATAGTTTCTGCTATGGCTAACAACCCTTCTGAGGGAC
CTTCTGATGATTTCTTTGACCAAATCTTGGGATTTCCTGCTTACAATGGAGCAGAACCTAATTTGG
CGGGAAATGATGCCGGAGCTATACCGCCGGCGATGATGCTCCAGCTTAACTCTGGTGATGGGTCAA
GTCAGTTTACTGGAGTGGGTCTTGGGGTTGGTTTAGGTGGTGGGGGGTTCCATGGTCATGGTGGGG
GTGGTTCTTTTCCTTTAGGGTTGAGTTTAGAACAAGGGAAAGGTGGGTTCTTGAAGATGGATGATG
TTTCAGCACCTGGAAGGAGGTTTAGAGATGATGTTGTTGATAGCAGAGCTTCTTCTTCTGTCAAAC
CTGGTTTTCATGGACAACCGATGCCTTCAATGCCGCATCCCCCTGCAATACGTCCAAGGGTGAGAG
CTAGGCGAGGACAAGCTACTGATCCACATAGCATAGCGGAGAGGTTACGTAGAGAGAGGATAGCAG
AAAGAATTAGAGCATTGCAAGAGTTGGTTCCCAGTGTCAATAAGACTGATAGAGCTGTAATGCTTG
ATGAAATTGTAGATTATATCAAATTCCTACGGCTGCAAGTGAAGGTGTTGAGCATGAGTAGGTTAG
GAGGAGCTGGTGCAGTAGCACCACTTGTTACAGACATTCCAATATCATCAGTAGAGGAAGAGAGCA
GTGAAGGTGGAAATAATAACCAACCAGCTTGGGAAAAGTGGTCAAGTGATGGCACAGAAAGGCAAG
TGGCTAAACTCATGGAAGAAAATGTTGGTGCTGCAATGCAATTTCTTCAGTCTAAAGCACTATGTA
TAATGCCTATTTCTCTTGCATCAGCAATTTATCACTCTCAGCCACCAGATACATCAAGTCTTGTTA
AGCCAGAAACAAATCCTCCTTCATAGATGAACACCCTTATAGAAAAAGGTGACATAACATGCTAAT
GACCTTAAAGAATTTTCTTGCAGTGGATGGTTCCTACTTTTATCTATTCAAGTAATATTTAGCTTT
TTGTTGAAAAAACAGGGTTCTTTTTTAAAAATATTTATGTATAAAGTAAACCATTTTCTATCATTT
ATATATTTGAGTAGTTTTGTTCT

SEQ ID NO: 383, S. lycopersicum TA21665

GCTCATTTTCCGTTCCTCTTAAAGCTCAATAAACCTATACGCCGGCAAACTCCGGCATGGCGGACA
ACCCACCGGAGGTATATGCTGCCGATGATTTTCTCGAGCAAATTCTGGCGATTCCCTCTTATGCTA
GCCTGCCGGTAACTGATTTAACCGCCGGTGCCTCATCGGAGAATTCAACATCTGGTGTCTCTCAGC
TCCAGCAGCAGCCGTTGTTCCCATTGGGGTTAAGTTTGGATAATGGCTTTGCTGACGCCAACAACA
CTGGAGGTTTTCAAGTGAAGACTGAAAGAGAAGCAATGAACATGGGGAATTTGTATCCAGGTCTAG
AACATTTGCAATCTCATGCTGTTTGCCTAAGTGTTCCTCAAGTTCACCAAGTCCAGCCTTTTCAAG
GCCACCCTACATCAAGCGCAATAGTAACAATACCACACCAACCTGCAATTCGTCCTAGGGTTCGGG
CACAAAGAGGACAGGCCACTGATCCACATAGTATTGCTGAGCGGTTAAGGAGAGAAAGGATATCAG
AACGAATAAAGGCTTTACAGGAACTTGTCCCCAGCTGCAATAAGACCGACAGGGCAGCAATGCTTG
ATGAGATTCTGGACTATGTGAAGTTCTTAAGGCTTCAAGTTAAGGTACTGAGCATGAGTAGGTTGG
GAGGAACTAGCGCGGCGGCACAAGTTGTTGCTGATATTCCATTGCAGTCTGTTGAGGGAGACACCT
GTGAAAGTCATTCCAACCAGCACGTCTGGGAGAAGTGGTCTGATTCTGAAACAGAGCAAGAGGTAG
CAAAGCTCATGGAGGAAGATGTTGGAACAGCCATGCAGTACCTTCAATCCAAATCACTCTGCATCA
TGCCTATTTCACTTGCTGCACTTATCTATCCGACTCAGCAAAGTGACAACCAATCAATGGTCAAGC
CAGAACAAGCGGCCCCATTGTAGACTTCTAAATTACTGCAGGTGCAATCAAGAACTCCTAATTTCT
TTTAGGCCCATCCGTCTATATATCCTTACCTATTTTCTCTTGAATCACCCTCTGAAATCCTTGGTA
ATGCAATCAAACTCAGATGACACAATTTTTGCTTTTAGCTCAATCATTGGCCCCTTTATCAACTTT
CTTGAGTTCAGTCACACCTTGAACAATGGCAACACTTAAGAAAGCAACCAATCCATCTATCTATAT
CATTGTAGGGAGGAGTGTATATTTTTTTTTCCAAACTCTCTGTTATTAGCTGCTTGTTTCCCCCTC
ACTCTTGTGGGTTGGTTGTGACTCCTTGTATCTTTTTTATAGTGGATAGCTAGCGATGTATTATAA
AAGAAAATTGTCCCAGCTTGCATTTGGTGGACTGTGCGGGGAGAAAGAAACCAAGTTTGTTTGAAG
ACAAAGCCATTAGTATACAGAAACTAAAGATGGAATGTTTAGGGCTTTTCTTTTTTTGGTGTAACC
ATAGCTTAATGGATGATTTAGAATCAACTTATGATGTACTAGATTCCTTGTGAGAAGATTAAGGAT
TAGGAGCTTCTTCATTTCTTTATTTCTGTTCTTTTGCTCATCTGGATGTATATATGGGGGATTACA
CACCCTTTGTGTAGTTTGTTATTAATACACAAAAATGTTACCTGTTCC

FIGURA 24 CONT.

SEQ ID NO: 384, S. lycopersicum TC172581

TTCCATTTCAGTTCCTCAAAACCCTTAACTTCTTCTCGCCGGGAAAAAAATTAATTTCGCCGGATA
AATCTCCGGCATGGCAGCAAACCAACCGGAAGGCTATGCCGACGATTTCCTCGAGCAAATTCTCGC
TATTCCTCCCTACTCTGGCTTGCCGGTTGCTGATGTTGGCACTCCATCAGAGACGACGTCGTTTAC
CTCGGCGTCTGCCGTATCTCATCTTAACTCTGCCGCTGCTGCTGGCCTACAGCAACCTTTGTTTCC
GCTGGGATTAAGCTTGGATAACGGCCGTGATGACGTCGGCGATGCAGGTCCTTATGCAGTGAAGCA
TGAAAGAGATGGAATGAATATCGGAAATCTATATGCAGGTCTAGAACACTTGCAATCTCATGCAGT
TCGTCACTCTGTGCCTTCTGTTCACCATGTCCAGCCTTTTCAAGGCCCACCAACAACGAGCACAAC
AGTAACTGTGCCACACCCACCTTCAATTCGTCCTAGGGTTCGAGCTCGGAGGGGACAAGCCACAGA
TCCACATAGCATTGCTGAGAGGCTGAGAAGAGAGAGGATATCAGAAAGAATAAAGGCTTTGCAGGA
ACTGGTACCCAGCTGCAATAAGACAGATAGGGCCGCAATGCTTGATGAGATTCTGGACTATGTGAA
GTTCTTAAGGCTTCAAGTTAAGGTACTGAGCATGAGTAGGCTGGGAGGAGCCAGTGCAGTGGCACA
ACTTGTTGCTGACATTCCATTACAATCTGTGGGAGGGGGACAGTGGTGAAAGTAGATCCAACCAGC
ATATATGGGATAAGTGGTCCAATGTTGACACAGAACGAGAGGTTGCTAAGCTCATGGAGGAAGACG
TCGGTGCAGCCATGCAATACCTTCAGTCTAAATCACTCTGCATCATGCCCATATCACTTGCTGCAC
TCATCTATCCTACTCAACAACCGGATGACCAGTCCCTGGTCAAGCCTGAAGCAGCAGCCCCATCAT
AGACCTTCAATTCGTTGCACGTGCCTCTAAGAACTCCCAATAGCCTTGTGTCCCCCCTCCTTTGTA
TCCTTACCTACCATCCATTTGTATTTTCCTGTGTAATTGGTGGCAATGCAATCAATGTCAGATGCC
ACAACTTTTGCTT

SEQ ID NO: 385, S. tuberosum TA30646

TTAGAAAGTAGCAACACACACAAACACCCAAAACACACACTTCCATTTCCATTTTTTTGCCTATAT
ATATATATATATATGTAACAAGTACTTGTGAGAGTACTAATTTGTACACCATCTTATCATTTTACTTG
TAAAGAACACTTTTTTAAGTAGTATTTGTTCAATTGAATGCTTTTCTGAGAAATTAAGGTGGTCGA
AAACAGCTCCGATCCAAGAAAGATTGAATCTTTATAGTTCTGCTATGGCTAACAACCCTTCTGAGG
GACCTTCTGATGATTTCTTTGACCAAATCATGGGATTTCCTGCTTACAATGGAGCAGAAACTAATT
TGGCGGGAAATGATGCCGGAGCTATACCGCCGGCGATGATGCTCCAGCTTAACTCCGGTGATGGGT
CGGGTCAATTTACTGGAGTGGGTCTTGGGGTTGGGTTAGGTGGTGGGGGGTTCCATGGTCATGGTG
GGGGTGCTTCTTTTCCTTTAGGGTTGAGTTTAGAACAAGGGAAAGGTGGGTTCTTGAAGATGGATG
ATGTTTCAGCACCTGGAAGGAGGTTTAGAGATGATGTTGTTGATAGCAGAGCTTCTTCTTCTGTCA
AACCTGGTTTTCATGGACAACCCATGCCTTCAATGCCGCATCCCCCTGCAATACGTCCAAGGGTGA
GAGCTAGGCGAGGACAAGCTACTGATCCACACAGCATTGCGGAGAGGTTACGTAGAGAGAGGATAG
CAGAAAGAATTAGAGCATTGCAAGAGTTGGTTCCCAGTGTCAATAAGACTGATAGAGCTGTAATGC
TCGATGAAATTGTAGATTATGTCAAATTCCTACGGCTGCAAGTGAAGGTGTTGAGCATGAGTAGGT
TAGGAGGAGCTGGTGCAGTAGCACCACTTGTTACAGACATTCCAATATCATCTGTAGAGGAAGAGA
GCAGTGAAGGTGGAAATAATAACCAACCAGCTTGGGAAAAGTGGTCAAGTGATGGCACAGAAAGGC
AAGTGGCTAAACTTATGGAAGAAAATGTTGGTGCTGCAATGCAATTTCTTCAGTCTAAGGCACTAT
GTATAATGCCTATTTCTCTTGCATCAGCAATTTATCACTCTCAACCACCAGATACATCAAGTCTTG
TTAAGCCAGAAACAAATCCTCCTTCATAGATGACACCCTTATAGAAAAAGGTGACATAACATGCTA
ATGACCTTAAGAATTTTCTTGAAGTGGATGGTTCCTACTTTTATCTATTTAAGTAATTTTTAGCTT
TTTGTTGAAAGAAACAGGGTTTTTTTTTTTTATGTTTATGTATAAATTAAACCATTTTCTATCATT
GATATATTCAAGTAGTTTTGTTTTCAGC

SEQ ID NO: 386, S. tuberosum TA28621

GATTTCAGTTCCTCAAAACCCTTAACTTCTTCTCGCCGGAAATTTTTTTTAATTTCGCCGGGTAAA
TCTCCGGCATGGCAGCAAACCAACCGGAAGCCTATGCCGATGATTTCCTCGAGCAAATTCTCGCTA
TTCCTTCCTACTCTGGCTTGCCGGTAGCTGATGTTGGCACTCCATCAGAGACGACGTCGTTTACCT
CGGCGTCTGCCGTATCTCATCTTAACTCTGCCGCTGCTGCTGGCCTACAGCAACCTTTGTTCCCGT

FIGURA 24 CONT.

TGGGATTGAGCTTGGATAACGGCCGTGATGACGTCAGCGAAGCTGGTGCTTATGCAGTGAAGCATG
AAAGAGATGGAATGAATATCGGAAATTTATATGCAGGTCTAGAACACTTGCAATCTCATGCAGTTC
GTCACTCTGTGCCTTCTATTCACCATGTCCAGCCTTTTCAAGGCCCACCAACAACGAGCACAACAG
TAACTGTACCTCACCCACCTTCAATTCGTCCTAGGGTTCGAGCTCGGAGGGGACAAGCCACAGATC
CACATAGCATTGCTGAGAGGCTGAGAAGAGAGAGGATATCTGAAAGAATAAAGGCTTTGCAGGAAC
TGGTACCCAGCTGCAATAAGACAGATAGGGCAGCAATGCTTGATGAGATTCTGGACTATGTGAAGT
TCTTAAGGCTTCAAGTTAAGGTACTGAGCATGAGTAGGCTGGGAGGAGCCAGTGCAGTGGCACAAC
TTGTTGCTGACATTCCATTGCAATCTGTGGAGGGGGACAGTGGTGAAAGTAGATCCAACCAGCATA
TATGGGATAAATGGTCCAATGTAGACACAGAACAAGAGGTTGCTAAGCTCATGGAGGAAGACGTCG
GTGCAGCCATGCAATACCTTCAGTCTAAATCACTCTGCATCATGCCCATATCACTTGCTGCACTCA
TCTATCCTACTCAGCAACCGGATGACCAGTCACTGGTCAAGCCTGAAGCGGCAGCCCCATCATAGA
CCTTCAATTCGTTGCACGTGCCTCTGAGAACTCCTAATAGCCTTGTGTCCCCCCTCCTTTGTATCC
TTACCTACCATCCATTTGTATTTTCCTGTGTAATTGGTGGCAATGCAATCAATGTCAGATGCCACA
ACTTTTGCTTTTATCCCAACCATTGGCCCTTTTTTTTTAATTACCTTTCTGGAGTTTCGTCAAACA
AGGCATGATGTCATCATACCTTGGAGAAAGGTAAACGTATGAAAGCACCCAATCAATTAGCATCAT
TTTCGGATGAGGTGAGCTTCTTTTTCAAAGTTATCTATGATTAGCTGCTCTTCTTCCACCATTTTG
TTGTGGGATTGTGGGGACTCCCTTTATCCTTTGCATAGTGGATGGTGGGGCTTTGTAAAAGTAACC
AACTGGTTGTTGGAAAGTATATTGTAATTTTAGTATATTATTGTAATGGGTAGTTGAAGTTGCTCT
CCCTCCAAAAAGTTTTG

SEQ ID NO: 387, S. tuberosum TA34455

CTTCTATCTTTCTCTCTTCTCCTTTTAGAACCAAACTTTCACACACAAAGCTACTGAAGAAAAAAT
TTCTCTCTGCAAAATCTCAGAGAGGAAAAAAAAAATGCAAGCCATGAATTCACTTCTAAGTCAACA
ACAGCAGTCACAGATGTCACTGCAAGACCTTCAAAATGGCGGAAATGGCGGTTCTACCGGTGGTGT
TGGTGGTTTAGGTCAACACAATATGGGTCACTTTCACTTTGATCCGACGTCGTCTCACGACGATTT
TCTCGAACAGATTCTCTCTTCTGTTCCTTCTTCTTCTCCTTGGCCTGACCTTTCCAAATCCTGGGA
CCCACATCACCATCTCTCATCGCCGCCGCATAACCCTAGCTCCGGCGAAGATCAACCTCCTTCTAA
TCCACTTCACTCACAGTTCCATTACGACGACCAGGCTTCTTCTCTACTAGCCTCAAAGCTCCGTCA
GCATCAGATCACTGGCGGCGGCGCGGCTGCAGCTGCTAAAGCACTTATGCTACAACAGCAGCTTTT
GCTTTCTAGAACACTCGCCGGAAACGGACTCAGGTCTCCTAATGGAGCTTCCGGCGATAACGGCCT
CCTTTCCCTACCTCTAAACCTCAGTAATAATGGTGACCAAAACGATGGCGTCGCTAACCCGGCTAA
TGACAATTCCGTTCAAGCTCTTTTCAATGGATTCACCGGATCTCTTGGTCAAACCTCCAATCAACC
TCAACATTTTCATCATCCTCAGGGAGGATCGATGCAATCGCAGAGT

SEQ ID NO: 388, S. tuberosum TA37666

GAAAGAAAAAAACAATGCAAGCTATGAACTCATTTCAAAGCACCGGCGAAAATGGTGCTTCGTCCG
GCGAACACATCAGTCATTCCCATTTCGATCCGTCGTCGTCACACGACGACTTCCTCCAACAGATTC
TCTCTTCCGTTCCTTCTTCTTCTCCCTGGCCTGAAATCTCCGGCGACGGTCACCCCTACAATTTCG
ACGACCATCAGTCAACTCTCTTGGCTTCCAAGCTCCGGCAGCATCAGATCAACGGCGGCAGCTCTG
CTGCCGCAGCTGCTAAAGCGTTGATGCTTCAGCAGCAGCTTCTGCTTTCTAGAGGAATCGCCGGTA
ATGGCGGCGATCAAAACGACGACGGTTTAAATCCGGGGAATGATATTTCAGTTCAAGCTCTTTACA
ATGGATTCGCTGGATCTCTTGGTCAAACCTCAAATCAATCTCAGCATTTTCACCATTCTCAGGCGC
AGAGTTTCGGAGCTCCGGCGGCGTCGTTGACGATGAATCAAACGCCGGCGGCGAGTGGTTCAGCTG
GTGGAGCGCAGCCAAAGCAACAGAAAGTTAGGGCTCGAAGAGGACAAGCAACTGATCCTCACAGCA
TTGCTGAAAGATTACGGAGAGAGAGAATTGCAGAGAGAATGAAGTCTTTGCAGGAGCTGGTACCTA
ATGCCAATAAGACAGACAAGGCTTCAATGTTAGATGAGATCATCGACTATGTCAGGTTCCTCCAGC
TCCAAGTCAAAGTTCTGAGTATGAGCAGATTGGGTGGTGCTGCAGCTGTTGCTCCCCTAGTTGCTG
ATAGATCCTCTGAGGGAGGAGGTGATTGTGTACAAGGAAATGGAGGTCGGGGTGGCAGTAATGGAA

FIGURA 24 CONT.

CGACGTCGTCTGCAAACAATGACAGCAGCATGACGATGACGGAGCACCAGGTAGCTAAGCTAATGG
AAGAAGATATGGGATCAGCAATGCAGTATCTTCAAGGAAAAGGCTTATCCCTTATGCCAATTTCCT
TAGCCACTGCTATTTCCACCTCCACGTGTCACTCCATGAAACCCAACAACCCACTACTACTCGGCG
GAGGCTCCGCTATCAACGGCGGTGGTGAGACCGGCGGTGGACCGTCTTCTCCTACCTTGTCAGCTT
CCACTGTTCAGTCAGCTACGATGGGCAATGGCGGGACTTGAGTTAGTCCTCCTATCCCAATGTGAC
TCACTGTTTTTATTTAGAATCAATTTGATTTTATATTTCTCGTTTTATACTATTATGGAATTATTT
ATAGTCATATTAATATTGATGACTTGTTTTTAGAAA

SEQ ID NO: 389, T. aestivum TC253044

CACGAGGGCAAGCCAAAGCCTCCGCCTCTCTCCTCCTTTTCCGCCCAAACAAACCGCCCCCTCCCC
TCCCGTCGCCCTCACCTCCGCCCTCGCCGGCAGTAGTATACGCCCGCCACCAGCTCCGACGACCCG
CGCCTTCCCCCCCCGCTCGCCCGCCGTCGTCTCCCCTGCCCGCTCCCCCGTCCCTCCGCCGCGGGTT
TGTTCTAGAGGGTGTATCTGAAGTTCTCTCAGCATCCGAGTGCGCAGCTTTGTACAAAAGGAACCA
TTCAGGCCATCTCAGATTGTCTTCATGGATGAATGATAGGATCACCAGTTGCTGTCACTTTAAGTA
TCTACAAGTCACTCTTTTCATAATATTTCTGAGCAGAGGTTGGGGAGTCAAGTTGTAAGGGTGTAA
AGATGGACTACTCTAATGGTTCTTTCTTTCCTTCATGGCCTGGCAATTCCGCTTCCGAGAATTATA
GCTTTGTTGATGGTTCAGTGGAATCATATGCAGAAGAAGGAAGTATGCCACCTACAGGCTATTTCA
GAGCTAGATCAAATCAGAATTTAACATTTGATGAGCATGAACAGAACCCTGCTATGCTTGCAAATG
GGTGCTTGCCGTACAACACCCAGACTGATCTATTATCTGGTGAGATTCTGTCAGAGGACAAACATT
CCAACAGCCTTATGGAGCTTTCCACAACTTCAGAACAATGGCAGTCTGCAAAGTAATTTAATCCCA
CCAGGGACTCTTCAGTGCACTTCAACACCTGGAACATTTGACCCGCAGTTGGATACCCCTGGCCTT
CTAGAACTTCCTCATGCCTTGTCCAGTTCAATTGAAAGCAATGGTAGTGAAGTTTCAGCTTTTCTT
GCTGATGTACATGCTGTTTCTTCAGCCTCAACTCTCTGTTCGACATTCCAAAATGTTCCTTCCTAC
ATGGAGCCAGTAAGCCTAGAAGCTTTCAGTTTTCAAGGGATACAAAATGCTCCTATGTTCAACAAT
ACAAGTCATTCAAATGGGAAACCTGTCAGTATTTGATGAGGCAACCATGGCATCACTACATGATAG
CAAAGAATTTCTCAGTGGTAGCATCTCATCTTTTGGTACGGCCGAGCAGTCACAACTAGCTGGTAG
TGGTTTGAAGGCTGAACAACAGGAACAAAATGCGATGTGCAATATTCCACTCCCTTTCGCTTCTGG
TAGTCAGATGGCAGTGAGTGAAGCACAAGGGGCAATGATTCCTTCAAAGATAAGCTCAACGATCCA
TAACAATAAAAGTGAGTACCCTGTCCCTATCAGCCATTCTGCTGATGCGCAGAACAAGGCAAATTC
AGCTAATGGAAACAGTGCCAGTGCTAAGCCACGAGCAAGGGCTCGTCGTGGACAGGCAACTGACCC
TCATAGTATTGCTGAACGGCTTCGCAGAGAGAAGATCTCAGAGAGGATGAAAAATCTCCAAGACCT
TGTACCAAACTCAAATAAGGCAGATAAATCATCAATGCTCGATGAAATAATTGATTATGTGAAATT
TCTTCAGCTTCAGGTGAAGGTCTTAAGCATGAGTAGGCTAGGAGCTCCCGGGGCAGTTCTTCCCCT
CCTTGCAGAATCTCAAACTGAGGGCCGTAGCAATTCACCTCTATCATCTCCAACCGCTTCACAAGG
GCTTCTGGATGCAGCAGGCCCAGAAGACAGCTTGGTCTTTGAGCAAGAAGTTATAAAGCTGATGGA
AACAAGCATCACAAATGCAATGCAGTACCTTCAGAACAAGGGCCTCTGCCTGATGCCTATCGCTCT
TGCTTCAGCCATATCCAACCAGAAAGGCACTTCTGCAGCTGCAATCCCTCCTGAAAGGTGAAAACA
GAAAGCACATCATGCTCCTCCCATCGGCCCGCGGAAACAACTGTCGAAATGTGCTAGAGTTCATAG
AAGTTGTGAGAAAACAAAAATCCGAGGCTGAAGGATAGGAAGTTATGCGACTAATAGTACAAGCTT
ATGGTATAATCTAGTAGCTTACTCCAAGAAACCTATACTTTTTCTCTTCTCTGTTGCACATCGGTT
GGTGTGTATTCTTTTGTTTTTTGGACGGCATTCTAACTGGGTACAAGTGTCTGGCACCTTGGGTTG
TACATGGATCTATGAGGGTTATTACCGACTGTTCATGATGTTATATGATGATCAAAGCAGACCTAG
GAAAACTCGGTTGCTTTGGGTTTCACTTGGTT

SEQ ID NO: 390, V. vinifera GSVIVT00016367001

ATGGATGAGTATCTGGATCACTTGTTTTCATCCACATCATGGTCAGATGTGAATGTGAAGGAGGGA
TCATCTTGGATTTGTGGTGAGCCTAGCCAAACAAATGGGATGCTTTCAGATTCAATTGGAATATAT
GAAGGTGATAAAAAAAACTCACCTGTTGGCATGACTAATTCAAACCTTATGATCGAGGGCTTGGTT

FIGURA 24 CONT.

ACTCAAGATACTTCATCTATTGTTCTTGGTGGAGAGTCTGATCATGGTCTTGGCAAGGGCTTACTT
TTGGAAGAAGCTCCACGGCAGCAAGACCTTCAAAACACTGAAGGCAATTCCTCCTTGAATGGAGCA
GTGAATGGCAGCTCAGAAGTTGGATACGTGGGCCTGCAACTAAATACTCCTACTTCAACCCCATGC
TCACTGGATCTTGGCTCTCCCAAGCAGCTTTCGCTGATTGGTGGCATGTCTAGGTCATCTCGTCCT
TTCACTGAGCTGGATCATGTTGGGTGCAATGGTAATGAGCCATCTGATTTTCAGAGATCAGTTGGA
AATTTTCAAGCATTGCCTCCAATTCCACAGTTGTGGTCTCAACCATCTTACGGGGGTGGTTCCTCC
TTGTCTCCTGTTATGGGAGAATACAAGATGCAGGGCTTTGGTCTGCAAGGAGAATATGTGGATAAT
GAAATGGATATCATGAGGAACAGATACGTTGGGGATGAAATTCTGCAACTTGATAACATTTCTTCA
GCAATCCCCATAAAGGGGAAAGAAGAGCAGCACACTCATCCTTTCTCTCCTTCTGCGGTTGGGCCC
CACATGACAATGACAGCATCTGGATTACAGTCTTTGCCACAGACTACGGTAGGTACTGCTTCTGGT
GGCTGTAATGGAACTGGAAAGCCGCGTGTAAGGGCTCGTCGAGGCCAGGCCACTGATCCTCACAGT
ATTGCTGAAAGGCTTCGGAGAGAGAAAATTGCTGAAAGGATGAAGAACCTGCAAGAACTTGTTCCC
AATTCCAATAAGACGGACAAAGCATCTATGCTCGACGAAATCATTGAGTACGTCAAATTTCTTCAG
CTCCAGGTCAAGGTACTAAGCATGAGTAGGCTGGGGGCAGCAGAGGCAGTTGTTCCTCTCATCACA
GATGGCCAAGCTGAGGGCTCTAAGGGTCTGTCACTTTCACCCTCGGCTGGTCAAGCTGAAGATATT
TGTCAATCTCCTGATCAAATTGCCTTTGAGCAAGAAGTAGTGAAGTTGATGGAGTCCAACGTGACC
ATGGCCATGCAATATCTTCAAAGCAAAGGTCTCTGCCTTATGCCAATTGCTCTTGCTACAGCCATA
TCCAGTGGAAAGGCAGCATCATCAGGTACTGGGTCCGATGAGGGAAAGAACGCAGCAACACCAGTA
GCAGCAACACCAGTAGCAACAGCTTACCTGGCATTGGGACACATCATACATCTTCTGATGGCAATG
TTCTGA

SEQ ID NO: 391, V. vinifera GSVIVT00017237001

ATGCTCTCTACTCTCCCTTCTTGGTCTGACCTCCCCGCAAACCCTAAATCTCCCTGGGAACTCAAT
GCTTCCAACCCTATCTCCATGCCTTCCAACAAGTCTAGGGATTTGTCCGACGATACCACGCCGTCT
AATCCGGACAACGTTCAGTTCGCTTTCGACGAGTCCGCCATGTTGGCCTCCAAGCTCCGCCAGCAC
CAGATCAGCGGGAACTCGTCGGCGGCCAAGTCGGCGTTGATGCTTCAGCAGCAGTTGCTTTTGTCA
CGAGGCGTCGCCATGGGTAGGTCCCCGTCGAATGGGTCCGGCGCCGGTGAGTCCGGCCTTCTTCAA
CTGCCTTTGTCACTTAGCAACGGGGATTCTTGCCTCGTTGACCGATCACAAAACGACGTCGTTGAT
GGGTCCTCCTCCTTCAAATCCCCCAATCAGGGAGGAGACGGTTCAGTTCAAGCTCTCTACAATGGC
TTCGCCGGAGCTCTTCACGGCTCCGGTCAAGCCTCCAACCAAGCTCAGAATTTCCACCATCCTCAG
GGTGGATCAATGCAGGCACAGAACTACGGAGCTCCGGCGACTAGGGTGAGGGCAAGAAGGGGTCAG
GCAACTGACCCACACAGCATCGCTGAGAGATTACGTAGAGAGAGAATTGCAGAGAGAATGAAGGCA
CTGCAGGAACTGGTTCCCAATGCCAACAAGACGGACAAGGCTTCAATGCTGGATGAGATCATTGAC
TATGTCAAATTCCTGCAGCTCCAAGTGAAGGTACTCAGCATGAGCAGGTTGGGCGGTGCCGCAGCT
GTTGCTCCCCTCGTTGCTGATATGTCCTCTGAGGGAGGTGGTGACTGTATACAGGCTAGTGGGACA
AGCGGCCCTACAGGAGGAAGGGCAACGAACGGAACACAAACAACCACATCAAACGACAGCCTCACA
GTGACGGAGCACCAGGTGGCCAAACTCATGGAGGAGGACATGGGTTCCGCCATGCAGTATCTTCAA
GGCAAAGGCTTATGCCTTATGCCCATCTCCCTCGCCACTGCCATCTCCACCACTACCTGCCACTCC
AGGAACCCCATGGTGGCTGCTGCCGCCGTCGCCGCCTCCAACATCAACAACGGCAGCCACACTCAC
CCACTCCTACCTAATTCCAACGCCGATGGCCCCTCCTCCCCCAGCATGTCCGTCCTGACCGTGCAG
TCAGCCACCATGGGGAACGGCCTGGCCGACGCGCCCGTAAAAGACGCTGCCTCCGTTTCCAAGCCC
TGA

SEQ ID NO: 392, V. vinifera TA46156

TCTTCCTTCTCACGGAAGCTGAGAGGCCGAGAGCTGAGACTTCACTCCAATCAAGGCCATCTCTTC
CATCTCCGAAACCCTAACCCACCTTCCGTACGGTTCCGACTCGTCCGAGTTTGAGTGATGGCGAGC
AATCCCTCGGAAGCTCCAGCGGACGATTTTCTCGAGCAAATCCTCGGAATTCCCACCTACCCCGCA
GCTGACCCTAATTTGGCTGCTAATGACGTGAATTTGGCCGCTCCTATGGTGCTTCAGCTCGGCTCC

FIGURA 24 CONT.

GGCGAAGGCTCTGGCCACATCGCCGGCGGGTACCAGGGAACCATGTTTCCGTTAGGGTTGAGGTTG
GAGCAGGGAAAGAGCAGCTTTCTGAAGCCCGAAGACGCGTCTGGCAGTGGCAAGCGCTTTCGTGAG
GAGGTTATTGATGGTAGAGCTTCTACCGTGAAAAATGCTTTCCATGGTCAACCAATGCAGGCTACC
GTTGCTGCAGCACCACATCCTCCTGCAATTCGCCCAAGGGTGCGGGCGAGACGTGGACAGGCCACA
GATCCACACAGCATTGCTGAGCGGTTACGCAGAGAAAGAATAGCAGAGAGAATTAGAGCACTGCAG
GAACTAGTTCCTAGTGTCAACAAGACGGATAGAGCTGCAATGCTTGATGAAATTGTGGATTATGTG
AAATTCCTAAGGCTCCAAGTAAAGGTCTTGAGCATGAGTAGATTGGGCGGAGCTGGTGCAGTTGCA
CCACTTGTAACAGACATTCCATTAGCATCAGTCGAGGAAGAAGCAAGTGAAGGGGGAAGAAATGAA
CCGGCATGGGAGAAATGGTCAAATGACGGTACAGAACGACAAGTTGCTAAGCTCATGGAAGAGAAT
GTTGGGGCTGCAATGCAGTTCCTTCAGTCCAAGGCACTCTGCATCATGCCCATCTCTCTCGCATCA
GCCATTTACCACTCTCAGCAGCCGGACACAACTCCACTGATCAAGCCCCAGACAAATCCCCCATCG
TAAAACCTCTCCCCACTCCC

SEQ ID NO: 393, V. vinifera TA46194

GGGAGCAAGGAAACCGAAGCAGGGCATTTCCCAAACATCTCTTACCGATTTATCGGACGACCTCTC
TGTGACCTCTTTCTCCAACAATTTGACCTTTTTGCTAAACCCGAAAACAATCCTCATCTCTTCTTT
GCTTCTCGCTGGAAAATCTATTTTCCGTTCCCGTCCATTCTCTAGCCATGGCCACCAACCCGCCCG
AAGGTTACGCCGATGATTTCCTGGAACAAATCCTCGCAATTCCCTCCTATCCTGGACACGACCCTA
ATATGGTGGGAACGGGTTCTACCATGGTGCTGCAGCTGAGTTCAGGCGACGGCTCAGGCCACGTCG
CCGGAGGTGGCTTCCAGGGGCCGGTCTTTCCGTTAGGTCTGAGCTTGGAGAGCGGGATTCCGGCGA
CACAGGTGGCGCCAGGAAGTGGCGAGCGGTTTCGAGATGACGTTGATGCCAGAGGTTCTTCAGGGA
GGAGCGAGAGAGAGTCGGTGCATTTGGATGGTTTGTTTCCGGCGTATGGACATGTACAGTCTCTCT
CCGTCCGACCCGCCGTTCCTCAAGTTCACCAGGCTTTTCATGGCCAGCCAACCCCTGTCCCAATTA
CTGCTGCGCCACACCCGCCAGCTATCCGCCCAAAGGTGCGTGCTCGGCGCGGACAAGCCACTGATC
CTCACAGCATTGCTGAGCGGCTACGTAGAGAGAGGATAGCAGAAAGAATGAAGGCTTTGCAAGAAC
TTGTCCCTAGCTCCAACAAGACGGATAGGGCTGCAATGCTTGATGAAATTGTGGACTATGTGAAGT
TCCTAAGGCTCCAAGTCAAGGTTTTGAGCATGAGTAGACTGGGAGGTGCTGGTGCAGTGGCACAAC
TCGTGGCTGACATCCCATTGCCAGCAGTTGAGGGAGAAACGGGTGAAGGTGGAAGCAACCAGCAAG
CCTGGGATAAATGGTCAAATGATGGCACAGAACGAGAAGTAGCAAAGCTCATGGAAGAAGATGTTG
GAGCTGCTATGCAGTTCCTCCAATCCAAAGCACTGTGCATCATGCCCATATCACTTGCCGCAGCAA
TATACCCTGCACACCAAACTGATACCCCCACGCTCATCAAGCCTGAGCCAAATGCTCCATCCTAGA
CCCCTCGAAAATGCACGTGCCTGTTGACAACTCCCATTGACCCCTAGTTGGCTAATCCCAAATTCC
TTAGTATCCTGAATCCATATCCTAAAGTGTTCCTCACTATCCCTTTTGGCAATGCAATCATAGTCA
GAAGCCATAGTCTTGCTTTTGCCTCTTTTTATTCTGAGTTTTGTCAAAGAAAGAGTGAATCATGCT
TTGGGAAAAAAGAAAAGGCACTTTCATCACTTCTTGCCATGCTCGAGTGAAGGGGGAAAATTTTCA
AAGCTATATTGATATGCTACTCCTTTTCATTTGTGCTGCTTGCTGGATGGCCAGAACACTTGTGGT
GGGGACTCTCTTTGGTTGATGCTGTGGTCTTGTAAAGCGGATGGGATGACTACGATTCCATGTAAT
TAAAATGAATTTGAAGAAAGCTGAAATGGATAATGGAAAGTCAATTTCAATGTTATATTGCATTAT
TTCCAGTCTTGGAGG

SEQ ID NO: 394, Z. mays TA139285

GTCGGCGGCAGCTGGAGAGCGAGGCCGAGCGACCAGTTACAGTTGCCGGCGGGCAAAGTGGGCCAC
TTGGTGCCAGGACATGGCCGGGCAACCGCCACCGCAGGGCCCCGAAGACGATTTCTTCGACCAGTT
CTTCTCCATGACGGCCGGCGGCTCCTACCCCGGCGCAACCGCGGGCGGCGGCCGCGCACCGGGTGA
CCAGCCGTTTTCCCTTGCGCTCAGCCTCGACGCCGCGGCCGCTGAGGCTTCCGGAAGCGGGAAGCA
CGCCGACGGTGGCAAGGCGGACCGGGAGGCTATACAGCTCCCTGGGCTCTTCCCGCCGGCGTTCGG
CGGCGGTGTGCAGCCACCCCACCTCCGCGCCACCCCGCCTACACAGGTGTTCCACGCGCAGCAGCC
GAAGCAAGGCGGTGCGGCAGTCGGGCCACAACCGCCGGCACCGAGGCCAAAGGTGCGGGCGCGGCG

FIGURA 24 CONT.

TGGGCAGGCGACCGACCCCCACAGCATCGCGGAGAGGCTAAGAAGAGAGAGAATAGCAGAAAGGAT
GAGGGCCCTACAGGAATTGGTGCCCAATACAAACAAGACAGATAGAGCAGCTATGCTAGATGAGAT
CCTAGATTATGTGAAGTTTCTGAGGCTTCAAGTCAAGGTTCTAAGCATGAGCAGGCTGGGTGGTGC
TGGTGCTGTGGCACAGCTGGTTGCTGATATCCCACTTTCAGTTAAGGGGGAAGCAAGTGATAGTGG
GAGCAAACAGCAGATTTGGGAGAAGTGGTCAACGGACGGCACAGAAAGACAGGTCGCAAAGCTGAT
GGAAGAAGACATTGGGGCAGCGATGCAATTCCTCCAGTCCAAAGCACTCTGCATGATGCCCATCTC
GCTCGCCATGGCAATCTATGACACGCAACATTCGCAGGACGGACAACCAGTAAAACCAGAACCCAA
CACTCCTTCCTAGTATAGTGGCAGCAAGCGTAACTTGCATCCTTTCGGGGCTGTTAGCAGGCGTTA
ACCTACCAATCAAAGCGTAAGATAGAAATGCCGGCTTCGATTCACTAACAAGGAGGAGGCAGGGAA
AAAAGACCTAAAAGATCCTCTCCACTTGAAGTTGTACCAAATGGTATATGATATCGCCTTTCCCTT
TCCGATCTTTTGTTCAAAGGTGCTTTAATTGCAAGGATATGGTAAGCAATAACCTGCTCATGCTCC
TCCATCTTTCAATGCCCCCTATGAGTGTGAGTGATGGAGAGAGGGAGCTGCTAAAAGAGTCTGGCA
TGCTCTCATAATGTTTCTTTCCGGGTCTGTAGTGCACTAGCTAGCCATACACTTTGTGGTGGGGGG
GCTGCTGCTTTTCTGGTAGTGGGTGTATATGGACTAATCAAAAGGGACTGATGTATGAGTGTGTCT
GCGGCCTGCGCAAATGGTTGACAAAGTTTAGATACAGGATGAAATGGAAGATGATGTTGATAGCGA
TGTTCTTCCTTCTCCATTTTTCCAAAGTGGAGGTCCGAGGCAGGCTTTGGTCAATGGAGGCTGAGC
GTGCGTTTCAGCGTTGCTAGTGGATCTGCTATCTGCGTTGAATACTTTATTTATATATGGATCCAC
ACTCTTGATAAGAACATTCG

SEQ ID NO: 395, Z. mays TA126400

GCTTCACACTTCCCCAGCAGAGCAGGCAAACCGCTCCCAGTCCCAGTCTCTCACCTGAGGTGACTG
GCGTCTCTTGTCGCAAGCCGCCAGAAGCTTCGGCCCCTGACGCGTGGGGGCCCCGTCCGCCCCCGC
CCCCGTCCCCACGTTGCCGGCCATGGCGGGGCAGCCACCGCCGTCCGGGGCCTCCGAGGACGACTT
CCTCGAGCACTTCTTCGCCTTCCCCTCCGCGGCCTCCGCCGGCGCTGCCGGGGGCCACGCGGGCGC
CGGTGTTGGCGGGGACCACCCCTTCCCCCTCGCCCTCAGCCTCGACGCCGCCGCCGAGGCCAAGCC
GGACCGTGATCCCGTGCAGCTCGCCGGCCTCTTCCCGCCGGTGTTCGCTGGCGCCGGCGGCGTGCA
CCAGCCGCACCTCCGCGGCCCACCGCCTCCGCAGATGTTCCAGGCGCAGCCGAAGCCGGGCGAGGG
AGGCATGGCGCCGCAGCCGCCAGCCCCGCGGCCCAAGGTGCGCGCGCGGCGGGGGCAGGCCACCGA
TCCCCACAGTATTGCGGAGAGGCTAAGGAGAGAGAGAATTGCAGAAAGGATGAGGGCATTGCAGGA
ATTAGTCCCCAACACAAACAAGACAGATAGGGCAGCCATGCTAGATGAGATCCTTGATTATGTGAA
GTTTCTTAGGCTTCAAGTAAAGGTTCTGAGTATGAGCAGACTGGGCGGTGCTGGCGCGGTTGCGCA
GCTGGTTGCTGATATTCCACTCTCAGTTAAGGGCGAGGCAGGCGACGGCGGGGGGGCGCCGCAGCA
GCAGCAGCAGCAGCACGTGTGGGAGAAGTGGTCGACGGACGGCACGGAGAAGCAGGTGGCGAAGCT
GATGGAGGAGGACATCGGGGCGGCAATGCAGTTCCTCCAGTCCAAGGCGCTGTGCATGATGCCGGT
CTCGCTCGCGATGGCGATTTACGACACCCAGCACCCCCTGGACGGCCACGGCCACTCGCTGAAGCC
CGAGCCCAACGCGTCATCCTAGTACAGTAACGTCGTCCTTAGCGTGCCTTCCAATTCCAACACCCA
CCCCCCTTCTCCCTAAAAAGTACTGTTAGGCGCCGATATTATCCGGAAGAAAAAAAGAAGTGTATG
ATATACCGAGCGTGCGTTTCCAGTGTGCTGACCCGTGGGATAGGAAAGGACCGAAAATTGTTGGTA
TATGATGATTATGTAGCCTCCTGTTTGCCTGTTCTGTTTCAAATCAAAGCAATGCTCATGCATCCA
GGTTTC

FIGURA 24 CONT.

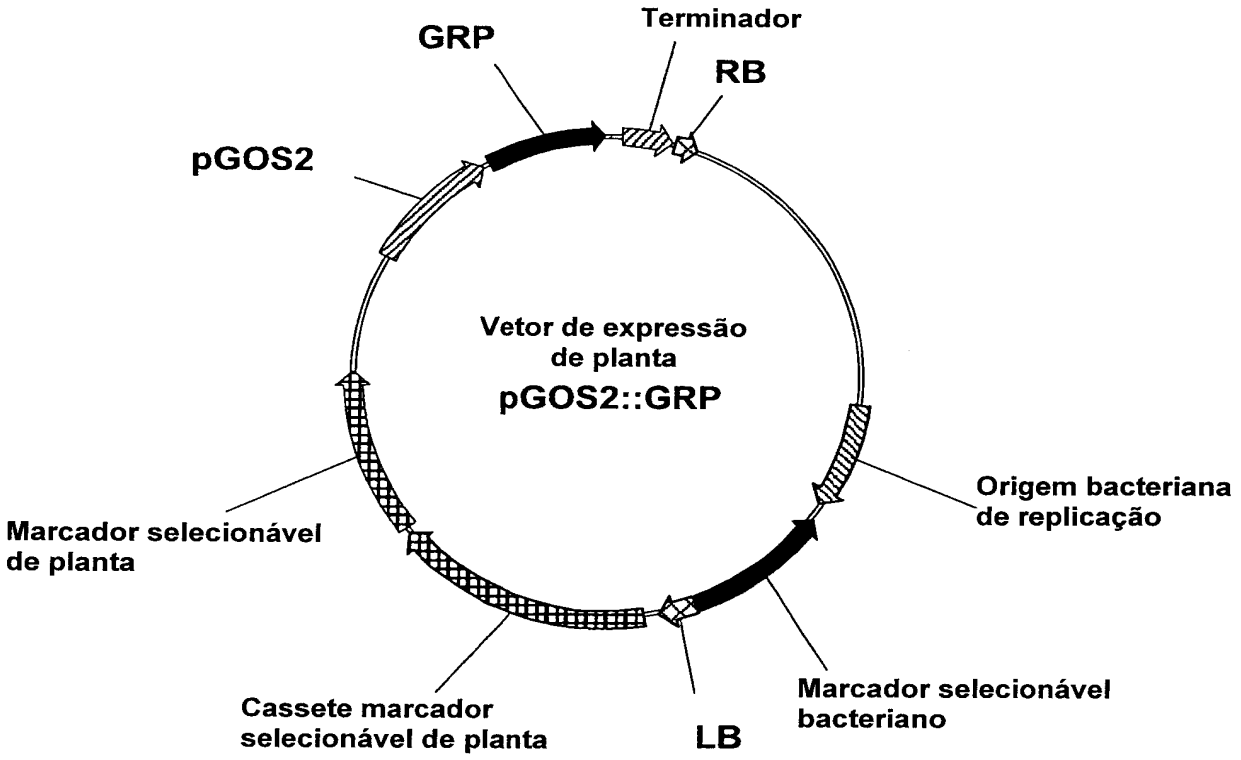

GRP
Terminador
RB
pGOS2
Vetor de expressão
de planta
pGOS2::GRP
Origem bacteriana
de replicação
Marcador selecionável
de planta
Marcador selecionável
bacteriano
Cassete marcador
selecionável de planta
LB

FIGURA 25

SEQ ID NO: 396, DNA - Oryza sativa

ATGGCGGAGGAGAAGCACCACCACCACCTGTTCCACCACAAGAAGGACGACGAGCCGGCCACCGGA
GTAGACTCCTACGGCGAGGGCGTCTACACGTCGGAGACGGTGACCACCGAGGTGGTCGCCGGCGGC
CAGGACGAGTACGAGAGGTACAAGAAGGAGGAGAAGCAGCACAAGCACAAGCAGCACCTCGGCGAG
GCCGGCGCCCTCGCCGCCGGCGCCTTCGCCCTGTATGAGAAGCACGAGGCGAAGAAGGACCCGGAG
AACGCGCACAGGCACAAGATCACGGAGGAGATCGCGGCCACGGCGGCGGTCGGCGCCGGCGGCTAC
GCCTTCCACGAGCACCACGAGAAGAAGAAGGACCACAAGAGCGCCGAGGAGTCCACCGGCGAGAAG
AAGCACCACCTCTTCGGCTGA

SEQ ID NO: 397, proteina - Oryza sativa

MAEEKHHHHLFHHKKDDEPATGVDSYGEGVYTSETVTTEVVAGGQDEYERYKKEEKQHKHKQHLGE
AGALAAGAFALYEKHEAKKDPENAHRHKITEEIAATAAVGAGGYAFHEHHEKKKDHKSAEESTGEK
KHHLFG

SEQ ID NO: 398, DNA - Oryza sativa - sequência de promotor GOS2

AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAATAT
AAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCACCTACT
TTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGTAATTAAGT
GGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGTTAAATTATTCG
AGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTTCTAGCTGAACTCAATGGGT
AAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTCTGAACGTATTGGCAAAGATTTAAA
CATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACATCATTAAGGACATGTCTTA
CTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATTTTTATTATTTATCTTTTTTCG
ATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAAT
ACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTAGGTAGCAATATC
TGAATTCAAGCACTCCACCATCACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTT
ACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAAAAAAGAATT
TTGCTCGTGCGCGAGCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCCACA
GAACAACCCACAAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAG
GCTTTGCGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGCATCCTCCTTCTCCCATCTATAA
ATTCCTCCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAAG
GACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTCGATCCATATCTTCCGGTCGAGTTCTTGGT
CGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCTCCCTTCGGTTGTTCTTGGATTT
ATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCC
TGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTCGGTTTGATTAGTAGT
ATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTAGGGATCGGAATCTTGCGATT
TTGTGAGTACCTTTTGTTTGAGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAGTAC
GGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTC
CCTATTGAACAAAAATAATCCAACTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTT
AAGCCTGTCCAAAATTTCGCAGCTGGCTTGTTTAGATACAGTAGTCCCCATCACGAAATTCATGGA
AACAGTTATAATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTT
TTTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATG
GTGCAAATCAGGTCTATATGATTGATTTTGGGCTGGCCAAGAAGTATAGAGACTCATCAACTCATC
AGCATATTCCGTATAGAGAAAACAAAAATTTGACAGGAACTGCTAGATACGCAAGCATGAATACTC
ATCTTGGCATTGAACAAAGTCGAAGGGATGATTTGGAATCGCTGGGTTATGTTTTAATGTACTTCT
TAAGAGGAAGTCTCCCTTGGCAGGGGCTGAAAGCAGGCACTAAGAAACAGAAGTATGAGAAGATCA
GTGAGAAGAAAGTATCAACATCAATAGAGACCTTGTGTAGGGGATATCCTGCAGAGTTTGCATCAT
ATTTTCATTACTGTCGATCACTAAGATTTGATGATAAACCAGATTATGCTTATCTGAAGAGAATTT

FIGURA 26

TCCGTGATCTTTTCATTCGTGAAGGGTTTCAATTTGATTATATATTTGACTGGACCATTTTGAAAT
ATCAGCAATCACAGCTTGCCAATCCTCCATCTCGTGCTCTTGGTGGTACTGCTGGGCCAAGCTCAG
GGATGCCTCATGCTCTTGTTAATGTTGAGAGGCAATCAGGTGGAGATGAAGGTCGACCAACTGGTT
GGTCTTCATCAAATCTTACACGTAATAAGAGCACGGGGCTGCATTTCAATTCTGGAAGCTTATTGA
AGCAAAAAGGCACAGTTGCTAATGATTTATCCATGGGTAAAGAGTTATCCAGTTCTAATTTTTTCC
GGTCAAGTGGACCATTGAGGCGTCCAGTTGTCTCTAGCATCCGAGACCCAGTGATTGCAGGGGGTG
AACCTGACCCCTCCGGCACTCTGACAAAAGATGCAAGCCCGGGACCATTGCGTAAAGTATCCAGTG
CTGCACGGAGGAGTTCACCAGTTGTGTCCTCAGATCACAAGCGCAGCTCCTCTATCAAAAATGCCA
ACATAAAGAATTTAGAGTCCACCGTCAAGGGAATAGAGGGTTTAAGTTTTCGATGATGAGGGACTG
CATTAGTAGCTGTGCTTTGTCTCAGTTCTCCGTTCACTGTAAATTTTGGCACACCAACTTGGGGAG
TAAGAGTTCTGATATTAGTTGCTGTCAGGAAGTACCATAAAGCTGAATTATACAATTAAAATTTGG
GATCCAATCGCAAAAGCACATTAAGGATATGATGGGGTTGCAGATCCAAACTCACAGATTCCAGTT
TATGCTCGTCCATACAGTTATAGGCACTTTCCATATTCTTTTCTTTAATCTCTGTCTCTTGCTTGT
TATTGTTATGTCGTGGTATTCTTGTTGAGGTCATGTTTGTGAATTGCGAAGATGGTCATGTATAAT
TGCCGAGAAATCATGTACTAGTTTGTTTTAAACATGAGCAAACTGTTATTTTGTTCAAGCTACTTT
AATATCAAAAAAAAAAAAAAAAGGGCGGCCGCTCTAGAGTATCCCTCGAGGGGCCCAAGCTTACGC
GTACCCAGCTTT

SEQ ID NO: 399, Sequência artificial **– prm06442**
GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGGCGGAGGAGAAGCA

SEQ ID NO: 400, Sequência artificial **– prm06443**
GGGGACCACTTTGTACAAGAAAGCTGGGTCGGCGACGTTGTGATGA

SEQ ID NO: 401, DNA - Zea Mays ZM06MC07637 58832771
ATGGCCGACTACTACCACGGCGGCAGGGCCATGTACTCCAACACTGACGAGTGCTATGACGCCGGC
AGGCACGGCGGCAGCGTGAGGGCGTACTCGCGCACCGACGAGTACTACGACAACGTGGACGACCGC
ATGAGGAGGCCCGCGTACGGCGCCGACGACTGCGGCTACGGCGGCAGGCAGACCGCCGTGTACGCC
GACGAGTACTCCCGCGGCGGCCGCTACGAAGGCTACGGCGTCGAGCAGGAGCACTTCAAGAGGGAG
GAGAAGGAGCATAAGCACAAGGAGCGCCTCGGCGAGGTGGGCGCCCTCGCGGGCGGCGCCTTTGCT
CTGTACGAGGGGCACCGTGCGAAGAAGGACCCGGCGAACGCGCAGAGGCACAAGATCGAGGCCGGC
GTGGCAACGGCGGCGGCGCTGGGCGCCGGCGGCTACGCGTACCACGAGCACCGCGAGCAGAAGGAG
GCCCACTACGAGAGCAGGCAGCAGGAGCACAGGGTGCCGCACGGGCACGGGCACGGGCACGGCTAC
TCCTACTACTGCGACTGA

SEQ ID NO: 402, proteina - Zea Mays ZM06MC07637 58832771
MADYYHGGRAMYSNTDECYDAGRHGGSVRAYSRTDEYYDNVDDRMRRPAYGADDCGYGGRQTAVYA
DEYSRGGRYEGYGVEQEHFKREEKEHKHKERLGEVGALAGGAFALYEGHRAKKDPANAQRHKIEAG
VATAAALGAGGYAYHEHREQKEAHYESRQQEHRVPHGHGHGHGYSYYCD

SEQ ID NO: 403, DNA - Zea Mays ZM06MC16908 TD146065794
TAGATAAATCGACGGTCGGTGCATATATATGAGCAAGACTGCCAGGCGCGCCCGGTAACAAATGAG
AATGGAAGCCGTCAGTAACTAGTGGTGATGGTGTCCACGGTGGCGGTGCTTCTTGGCGTCTTTCTT
CTGGTGATGCTCGTGGAAGGCAAAGCCGGCGCTGCCCACGGCGGCGGCGGCGGCGACCCCTTCCTT
GATCTTGTGCGAGTGCGCGTGCTCCGGGTCCTTCTTCGCCTTCTGCTTCTCGTGCATAGCGTAGGC
TCCTGCAGCGACGGCGCCGAGCTGGCCGAGCTGCTCCATGTGCTTGTGGTGCTTCTCCTCCTTGCG
CCAGTCGGTCACCTTGGCCGCCTCTCCTCCATGGTGCTTGTTGCCCAGGAAGTGGTGCGCCATTGT
TTCAGTCGTTGGAGATGCGCTAGTCGACGGCCGGTTGTGTCTTGGGGTTCGTTCGATCTGCTTGAT
GATGAT

FIGURA 26 CONT.

SEQ ID NO: 404, proteina - Zea Mays ZM06MC16908 TD146065794
MAHHFLGNKHHGGEAAKVTDWRKEEKHHKHMEQLGQLGAVAAGAYAMHEKQKAKKDPEHAHSHKIK
EGVAAAAAVGSAGFAFHEHHQKKDAKKHRHRGHHHH

SEQ ID NO: 405, DNA - Zea Mays ZM06MC35104 62124878
ATGGAGCAGCTCGGCCAGCTCGGCGCCATCGCCGCCGGCGCGTACGCTCTGGTGAGACCAAAAAAA
CCCAACCTCCATCCCTTTGGTCGCCTGATCGATGCATGCCATGGTCCAGCTCTTGATCTGTGTGTG
TCTGTGCGTGTGCATCGACGTGTGGCTGCAGCACGAGAAGCACAAGGCCAAGAAGGACCCGGAGAA
CGAGCACGGCCACCGGGTCAAGGAGGAGGTGGCCGCCGTCGCCGCCGTGGGCTCCGCCGGGTTCGC
CTTCCACGAGCATCACGAGAAGAAGGACGCCAAGAAGCACGCCCACAACTGATCCGTCGCAGGCTG
CCTCAGCCGAAGCAGCAGCACATCGTCGCGGTTGCTGTTTCATCTGTTCTCCCAGCCTCGTCTTCG
TCTACTGTGTGCCGGGCAGGCCTTGATTTGAGGCTTTG

SEQ ID NO: 406, proteina - Zea Mays ZM06MC35104 62124878
MEQLGQLGAIAAGAYALVRPKKPNLHPFGRLIDACHGPALDLCVSVRVHRRVAAAREAQGQEGPGE
RARPPGQGGGGRRRRRGLRRVRLPRASREEGRQEARPQLIRRRLPQPKQQHIVAVAVSSVLPASSS
STVCRAGLDLRL

SEQ ID NO: 407, DNA - Zea Mays ZM06MC38573 59496278
ATGGCGGAGGAGAAGCACCACCACCACCACCTGTTCCACCACAAGAAGGACGAGGAGCAGCTCGCC
GCCGGCGGGTACGGCGAGTCCGCCGAGTACACGGAGGCCACGGTGACGGAGGTGGTGTCCACGGGC
GAGAACGAGTACGACGAGTACAAGAAGGAGGAGAAGGAGCACAAGCACAAGCAGCACCTCGGCGAG
GCCGGCGCCATCGCCGCCGGCGCCTTCGCACTCTACGAGAAGCACGAGGCGAAGAAGGACCCGGAG
CACGCGCACCGCCACAAGATCGAGGAGGAGGTCGCGGCGGCGGCGGCCGTCGGCTCCGGCGGCTTC
GCCTTCCACGAGCACCACGAGAAGAAGAAGGACCACAAGGAGGCCGAGGAGGCCGGCGGCGAGAAG
AAGCACCACTTCTTCGGCTGA

SEQ ID NO: 408, proteina - Zea Mays ZM06MC38573 59496278
MAEEKHHHHHLFHHKKDEEQLAAGGYGESAEYTEATVTEVVSTGENEYDEYKKEEKEHKHKQHLGE
AGAIAAGAFALYEKHEAKKDPEHAHRHKIEEEVAAAAAVGSGGFAFHEHHEKKKDHKEAEEAGGEK
KHHFFG

SEQ ID NO: 409, DNA - Zea Mays ZM06MC39332 62040304
atgtcggaggagaagcaccaccacttcttcaaccaccacaagaaggacgaggagcagcccgccggc
gagtacgggtactccgagaccgaggtggtcaccgccaccggcgagggcgagtacgagaggtacaag
aaggaggagaaggagcacaagcacaagcagcacctcggcgaggccggcgccatcgccgccggcgcc
ttcgcactctacgagaagcacgaggcgaagaaggaccccgagcacgcgcacaggcacaagatcacg
gaggaggtcgctgccgcggcggccgtcggcgccggcggctacgtattccacgagcaccacgagaag
aagaaggaccacaaggacgccgaggaggccagcggcgagaagaagcaccaccacctcttcggctga

SEQ ID NO: 410, proteina - Zea Mays ZM06MC39332 62040304
MSEEKHHHFFNHHKKDEEQPAGEYGYSETEVVTATGEGEYERYKKEEKEHKHKQHLGEAGAIAAGA
FALYEKHEAKKDPEHAHRHKITEEVAAAAAVGAGGYVFHEHHEKKKDHKDAEEASGEKKHHHLF

SEQ ID NO: 411, DNA - Zea Mays ZM06MC39358 61988082
ATGGAGAAGCTCGGCGAGCTCGGCGCCATCGCCGCCGGCGCGTACGCCCTGCACGAGAAGCACAAG
GCCAAGAAGGACCCAGAGAACGAGCACGGGCACCGGGTCAAGGAGGAGGTGGCCGCCGTCGCCGCC
GTGGGCTCCGCCGGCTTCGCTTTCCACGAGCACCACGAGAAGAAGGACGCCAAGAAGCACGCCCAC
AACTGA

FIGURA 26 CONT.

SEQ ID NO: 412, proteina - Zea Mays ZM06MC39358 61988082 |ZM06MC39358 61988082|MAIZE

MEKLGELGAIAAGAYALHEKHKAKKDPENEHGHRVKEEVAAVAAVGSAGFAFHEHHEKKDAKKHAH
N

SEQ ID NO: 413, DNA - Zea Mays ZM06MC42566 TD99014611

TTGTCACTTGCTCTCCCTCCAACAAGCTAATTAAGGCCGGTCATCCCTCTTCTAGCTCGTTTCATT
ATCCATGGCGGAGGAGAAGCACCACCACCACCACCTGTTCCACCACAAGAAGGACGAGGAGCAGCT
CGCCGCCGGCGGGTACGGCGAGTCCGCCGAGTACACGGAGGCCACGGTGACGGAGGTGGTGTCCAC
GGGCGAGAACGAGTACGACGAGTACAAGAAGGAGGAGAAGGAGCACAAGCACAAGCAGCACCTCGG
CGAGGCCGGCGCCATCGCCGCCGGCGCCTTCGCACTCGTACGTAGTCCGATCGATCCGATCCTCCT
TGAGTAGTATATATATATACATACATGAACGAGAAAGAATAATATATATATTAATCGAACGAACTG
AATGACGGTCACCTCGTGTGACGTGGACATGCACAGTACGAGAAGCACGAGGCGAAGAAGGACCCG
GAGCACGCGCACCGCCACAAGATCGAGGAGGAGGTCGCGGCGGCGGCGGCCGTCGGCTCCGGCGGC
TTCGCCTTCCACGAGCACCACGAGAAGAAGAAGGACCACAAGGAGGCCGAGGAGGCCGGCGGCGAG
AAGAAGCACCACTTCTTCGGCTGATTGATCCTCCCGTATCGTCGTCCCTCCCCGTGTGCTACGCGT
GCGTGTGAGAGTGATATCGAGCGCCCGCCGTGTTGTGCGCGCGTACGTATGTATGCGCTCGTGTGA
TGCACGAATAAGCGTGGCTACGTAATCTATCGTATGTATACGTGTGTGTATGCATGTGCTTGTGTA
TGATCGTGGTACGAGGACCGAAAAAATGTATGCAACTCTGATTTACTTACATGTTTAGTTGTTTAC
GTTTCTCCAAAAGTATTAGTACCCATTATGTTATATTTTACTATATTATATTTTATTGATTATTAA
ACTATGTATGGATGAAATTACTATTTCAAGCTGGTAGCTAGACCGAGATAAAACTCGTCCGTGTTT
TTTAACGGGTGTATTATT

SEQ ID NO: 414, proteina - Zea Mays ZM06MC42566 TD99014611

MAEEKHHHHHLFHHKKDEEQLAAGGYGESAEYTEATVTEVVSTGENEYDEYKKEEKEHKHKQHLGE
AGAIAAGAFALVRSPIDPILLE

SEQ ID NO: 415, DNA - Zea Mays ZM06MS58275860.r021

AGAAGGAGCATAAGCACAAGGAGCGCCTCGGCGAGGTGGGCGCCCTCGCGGGCGGCGCCTTTGCTC
TGGTAATATATATAATACAGTTACCATATATATATGTGCATGCACGATCTATATATGTGGTGATGT
GTTGTTGATGGAATATAATATGTGGAAATTAGTACGAGGGGCACCGTGCGAAGAAGGACCCGGCGA
ACGCGCAGAGGCACAAGATCGAGGCCGGCGTGGCAACGGCGGCGGCGCTGGGCGCCGGCGGCTACG
CGTACCACGAGCACCGCGAGCAGAAGGAGGCCCACTACGAGAGCAGGCAGCAGGAGCACAGGGTGC
CGCACGGGCACGGGCACGGGCACGGCTACTCCTACTACTGCGACTGAAACGACTGATCATCCCATC
GATCGAGATATGCATGCAGAACTCGATATATATACGGCCCCTAGCGCCGTACACGTGCGTGCGTGT
AAGTAAATAAGAGCAAACGTCGTCACGTGTGTGTTGCTCGATCACTACTCCATCGATCTACTTATT
ATCGTCTGCAAATAACGACTAAGCGCATGCTTCTGCAGCGTGTAACGATCAAAAAAAAAA

SEQ ID NO: 416, proteina - Zea Mays ZM06MS58275860.r021

MEYNMWKLVRGAPCEEGPGERAEAQDRGRRGNGGGAGRRRLRVPRAPRAEGGPLREQAAGAQGAAR
ARARARLLLLLRLKRLIIPSIEICMQNSIYIRPLAPYTCVRVSK

SEQ ID NO: 417, DNA - Zea Mays ZM06MC07637 58832771

GGTCGACGATTTCGTCGGCCCAACACTAGTAGCGTAGCTGAGAGCCTGAGACGATGGCCGACTACT
ACCACGGCGGCAGGGCCATGTACTCCAACACTGACGAGTGCTATGACGCCGGCAGGCACGGCGGCA
GCGTGAGGGCGTACTCGCGCACCGACGAGTACTACGACAACGTGGACGACCGCATGAGGAGGCCCG
CGTACGGCGCCGACGACTGCGGCTACGGCGGCAGGCAGACCGCCGTGTACGCCGACGAGTACTCCC
GCGGCGGCCGCTACGAAGGCTACGGCGTCGAGCAGGAGCACTTCAAGAGGGAGGAGAAGGAGCATA

FIGURA 26 CONT.

AGCACAAGGAGCGCCTCGGCGAGGTGGGCGCCCTCGCGGGCGGCGCCTTTGCTCTGTACGAGGGGC
ACCGTGCGAAGAAGGACCCGGCGAACGCGCAGAGGCACAAGATCGAGGCCGGCGTGGCAACGGCGG
CGGCGCTGGGCGCCGGCGGCTACGCGTACCACGAGCACCGCGAGCAGAAGGAGGCCCACTACGAGA
GCAGGCAGCAGGAGCACAGGGTGCCGCACGGGCACGGGCACGGGCACGGCTACTCCTACTACTGCG
ACTGAAACGACTGATCATCCCATCGATCGAGATATGCATGCAGAACTCGATATATATACGGCCCCT
AGCGCCGTACACGTGCGTGCGTGTAAGTAAATAAGAGCAAACGTCGTCACGTGTGTGTTGCTCGAT
CACTACTCCATCGATCTACTTATTATCGTCTGCAAATAACGACTAAGCGCATGCTTCTGCAGCGTG
TAACGATCCATCACTGTTACTATATCTAATGATGCATGCATGCAATTAAGCTATATATGTAAAAAA
CTATGCGTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANCCAACCTAAA
AAAAAAAAAAAAAA

SEQ ID NO: 418, proteina - Zea Mays ZM06MC07637 58832771
MADYYHGGRAMYSNTDECYDAGRHGGSVRAYSRTDEYYDNVDDRMRRPAYGADDCGYGGRQTAVYA
DEYSRGGRYEGYGVEQEHFKREEKEHKHKERLGEVGALAGGAFALYEGHRAKKDPANAQRHKIEAG
VATAAALGAGGYAYHEHREQKEAHYESRQQEHRVPHGHGHGHGYSYYCD

SEQ ID NO: 419, DNA - Lycopersicon esculentum
ATGGAGGAGGAGAAACACCACCACCACCACCTGTTCCACCACAAGGACAAGGCGGAGGAGGGCCCC
GTCGACTACGAAAAAGAAATCAAACACCATAAACATCTCGAGCAAATCGGTAAACTTGGCACTGTT
GCTGCCGGTGCCTACGCCTTGCATGAGAAACATGAGGCAAAGAAAGATCCAGAACATGCACACAAA
CACAAGATAGAGGAAGAGATAGCAGCAGCTGCTGCAGTTGGGGCAGGTGGATTTGCATTCCATGAG
CATCATGAGAAAAAAGATGCCAAGAAAGAAGAAAAAAAAAAGCTGAGGGGGGACACCACCATCTCT
TCTAAATTGTTATTTTAG

SEQ ID NO: 420, proteina - Lycopersicon esculentum
MEEEKHHHHHLFHHKDKAEEGPVDYEKEIKHHKHLEQIGKLGTVAAGAYALHEKHEAKKDPEHAHK
HKIEEEIAAAAAVGAGGFAFHEHHEKKDAKKEEKKKLRGDTTISSKLLF

SEQ ID NO: 421, DNA - ASR Lilium longiflorum
ATGGCCGAGGAGCACCACAAGCACCACCTCTTCCATCACAACAAGGAGAGCAACCCGGAGGTGGTC
GTCTCCGAGACCAACTACATCACCGACGGCTTAAACTCCACCTACAGCACATCCAGCGATGTATAC
GGTGGCAACCAGTCACAGGCCAACTACGAGGACTACGAGAAAGAGAAGAAGCACATCAAGCATAAG
GAGCATCTCGGCGAGATGGCCACGGCCGGCGCCGGTGCTTTTGCCCTTTACGAGAAGCACGAGGCA
AAGAAGGACCCCGAGCACGCCCACAGGCACAAATTGGAGGAGGAGATTGCCGCAGCTGCAGCTGCG
GGGGCTGGAGGGTACACCTTCCATGAGCACCACGAGAAGAAAACCTTGAAGAAGGAGAATGAGGAA
GTCGAGGGGAAGAAGCACCACTTCTTCGGTTAA

SEQ ID NO: 422, proteina - ASR Lilium longiflorum
MAEEHHKHHLFHHNKESNPEVVVSETNYITDGLNSTYSTSSDVYGGNQSQANYEDYEKEKKHIKHK
EHLGEMATAGAGAFALYEKHEAKKDPEHAHRHKLEEEIAAAAAAGAGGYTFHEHHEKKTLKKENEE
VEGKKHHFFG

SEQ ID NO: 423, DNA - gil15667622:130-558 Saccharum officinarum SoDip22 mRNA para proteína de 22 kD indutível por secura, cds completos
ATGGCGGAGGAGAAGCACCACCACCACCTGTTCCACCACAAGAAGGACGAGGAGGAGCAGGTGGAG
CAGCCCGCCGGCGGCGGCGGGTACGGCGAGGCCGCCGAGTACACGGAGACCACGGTGACGGAGGTG
GTGTCCACGGGCGAGGACGAGTACGACAAGTACAAGAAGGAGGAGAAGGAGCACAAGCACAAGCAG
CACCTCGGCGAGGCCGGCGCCATCGCCGCCGGCGCCTTCGCACTCTACGAGAAGCACGAGGCGAAG

FIGURA 26 CONT.

AAGGACCCGGAGCACGCGCACCGCCACAAGATCGAGGAGGAGGTCGCGGCCGCGGCGGCCGTGGGA
TCCGGCGGCTTCGCGTTCCACGAGCACCACGAGAAGAAGAAGGACCACAAGGAGGCCGAGGAGGCC
GGCGGCGAGAAGAAGCACCACTTCTTCGGCTGA

SEQ ID NO: 424, proteina - gil15667622:130-558 Saccharum officinarum SoDip22 mRNApara proteína de 22 kD indutível por secura, cds completos

MAEEKHHHHLFHHKKDEEEQVEQPAGGGGYGEAAEYTETTVTEVVSTGEDEYDKYKKEEKEHKHKQ
HLGEAGAIAAGAFALYEKHEAKKDPEHAHRHKIEEEVAAAAAVGSGGFAFHEHHEKKKDHKEAEEA
GGEKKHHFFG

SEQ ID NO: 425, DNA - gi|17128027|emb|AX297905.1| Sequência de DNA da Patente WO0183756

TGTCGATCCAATTGTCACTTGCTCTCCCTCCAACAAGCTAATTAAGGCCGGTCGTCATCCCTCTTC
TAGCTCGTTTTATTATCCATGGCGGAGGAGAAGCACCACCACCACCACCTGTTCCACCATAAGAAG
GACGAGGAGCAGGAGGAGCAGCTCGCCGGCGGCGGGTACGGCGAGTCCGCCGAGTACACGGAGGCC
ACGGTGACGGAGGTGGTGTCCACGGGCGAGAACGAGTACGACGAGTACAAGGAGGAGAAGCAGCAT
AAGCACAAGCAGCACCTCGGCGAGGCCGGCGCCATCGCCGCCGGCGCCTTCGCACTCTACGAGAAG
CACGAGGCAAAGAAGGACCCGGAGCACGCGCACCGCCACAAGATCGAGGAGGAGGTCGCGGCGGCG
GCGGCCGTCGGCTCCGGCGGCTTCGCCTTCCACGAGCACCACGAGAAGAAGAAGGACCACAAGGAC
GCCGAGGAGGCCGGCGGCGAGAAGAAGCACCACTTCTTCGGCTGATTGATCCCTCCCGTATCGTCG
TCCCTCCCCGTGTGCTACGCGTGCGTGTGTGAGAGTGATATCGAGCGCCCGCCGTGTTGTGCGCGC
GTACGTATGTATGCGCTCGTGTGATGCACGAATAAGCGTGGCTACGTAATCTATCGTATGTATACG
TGTGTGTATGCATGTGCTTGTGTATGATCGTGGTACGAGGACCGAAAAAATGTATGCAACTCTGAT
TTACTTACATGTTTAGTTGTTTACGTTTCTCCAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 426, protein - gi|17128027|emb|AX297905.1| Sequência de DNA 1 da Patente WO0183756

MAEEKHHHHHLFHHKKDEEQEEQLAGGGYGESAEYTEATVTEVVSTGENEYDEYKEEKQHKHKQHL
GEAGAIAAGAFALYEKHEAKKDPEHAHRHKIEEEVAAAAAVGSGGFAFHEHHEKKKDHKDAEEAGG
EKKHHFFG

FIGURA 26 CONT.

MDCNMVSSSQ**WDWEH**LIMSNPSRTEDDSKQLPTEWEIEKGEGIESIVPH
FSGLERVSSGSATSFWHTAV***SKSSQSTSINSS***SPEAKRCKLASESSPGD
SCSNIDFVQVKAPTALEVSVASAESDLC***LKLGKR***TYSEEYWGRNNNEIS
AVSMKLLTPSVVAGKSKLCGQSMPVP*RCQIDGCELDLSSAKGYHRKHKV*
CEKHSKCPKVSVSGLERRFCQQCSRFHAVSEFDEKKRSCRKRLSHHNAR
*RRKPQGVFSMNPERVYDRRQHTNMLWNGV*SLNARSEEMYEWGNNTYDTK
PRQTEKSFTLSFQRGNGSEDQLVASSSRMFSTSQTSGGFPAGKSKFQLH
GEDVGEYSGVLHESQDIHR***ALSLLS***TSSDPLAQPHVQPFSLLCSYDVVP
K

FIGURA 27

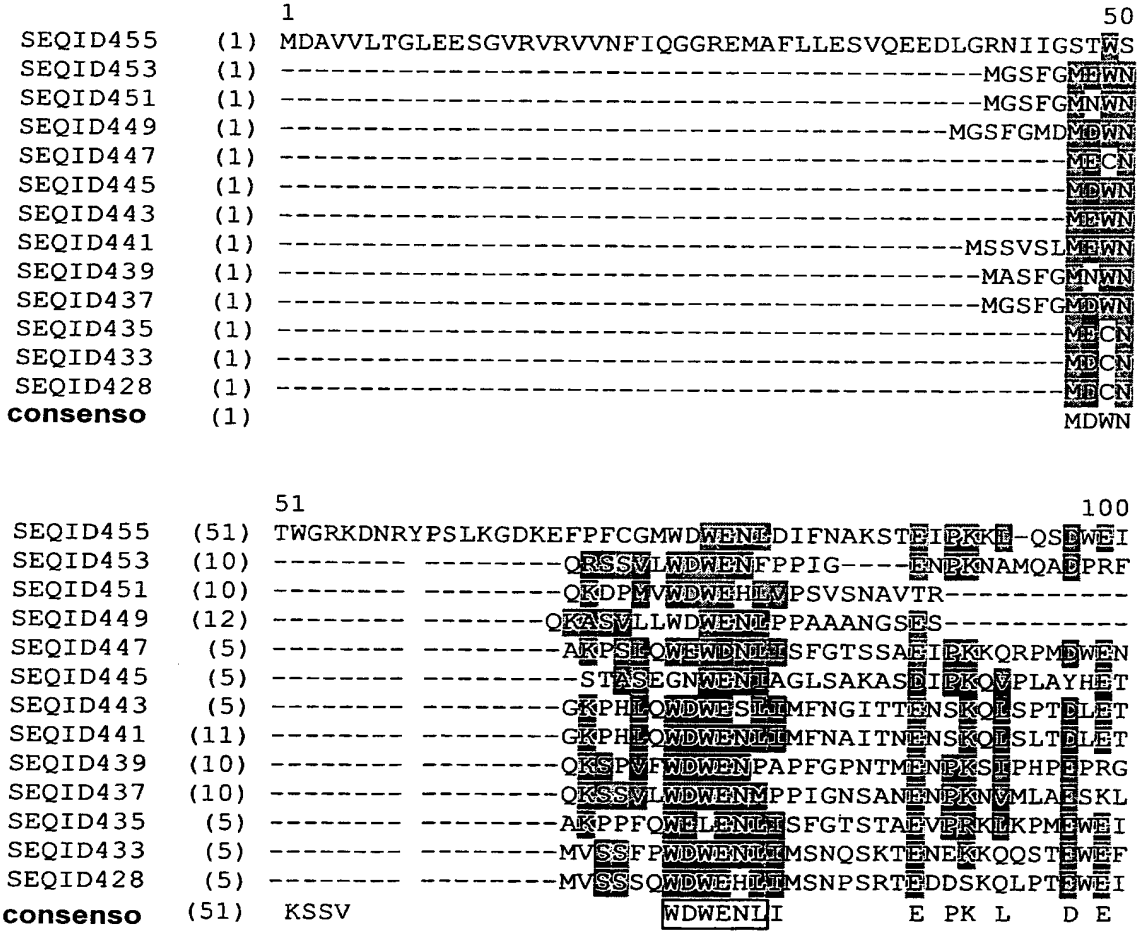


FIGURA 28

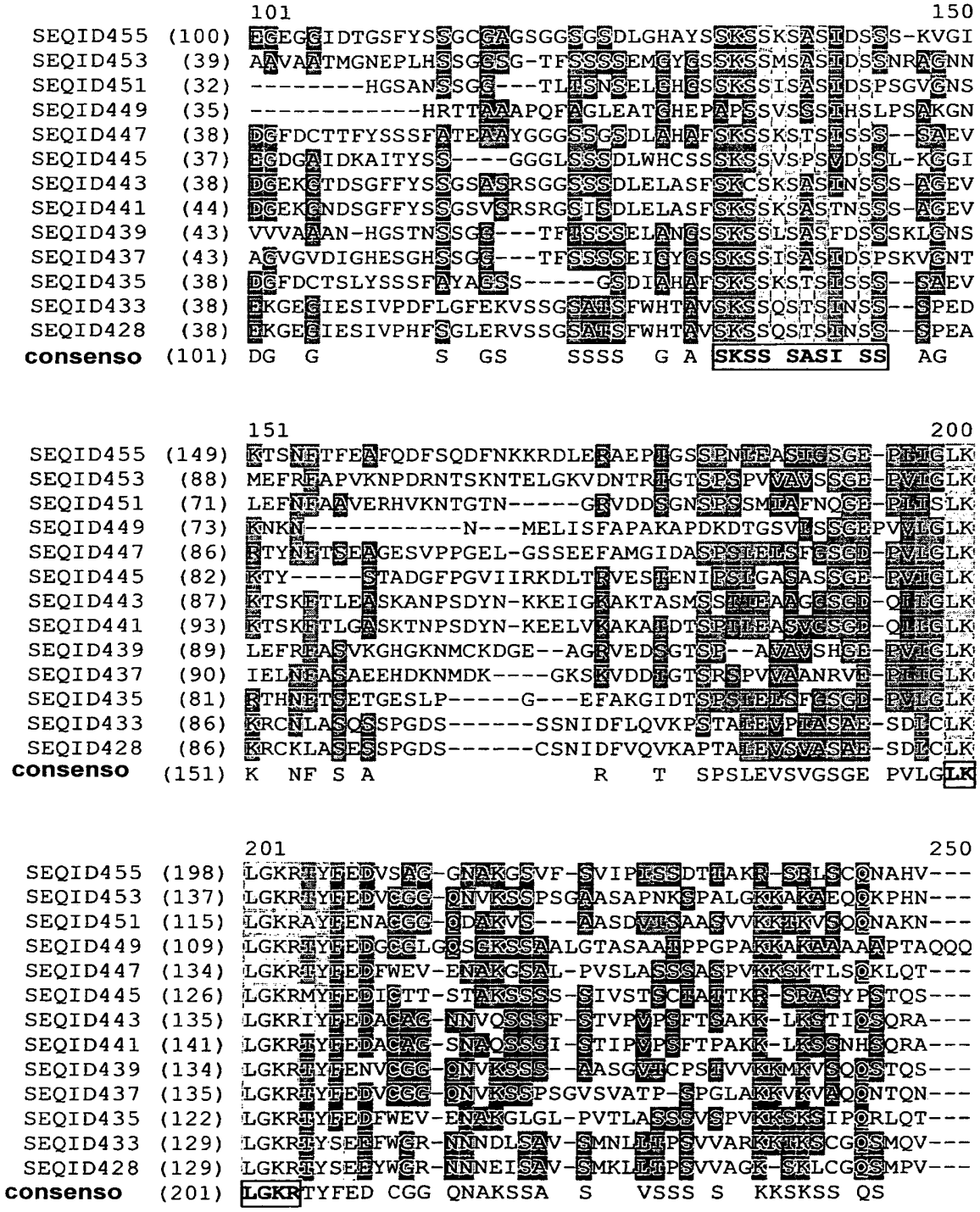


FIGURA 28 CONT

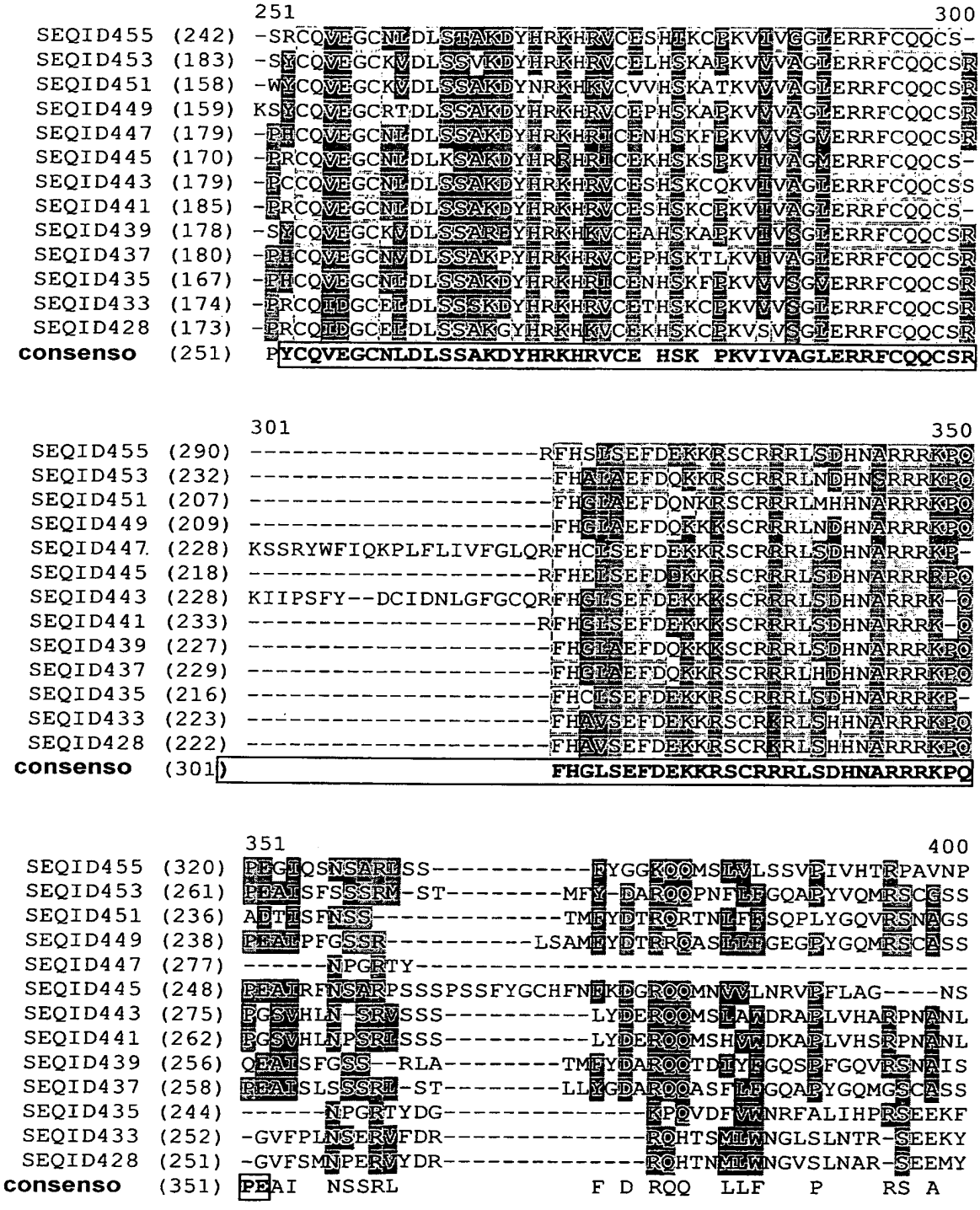


FIGURA 28 CONT

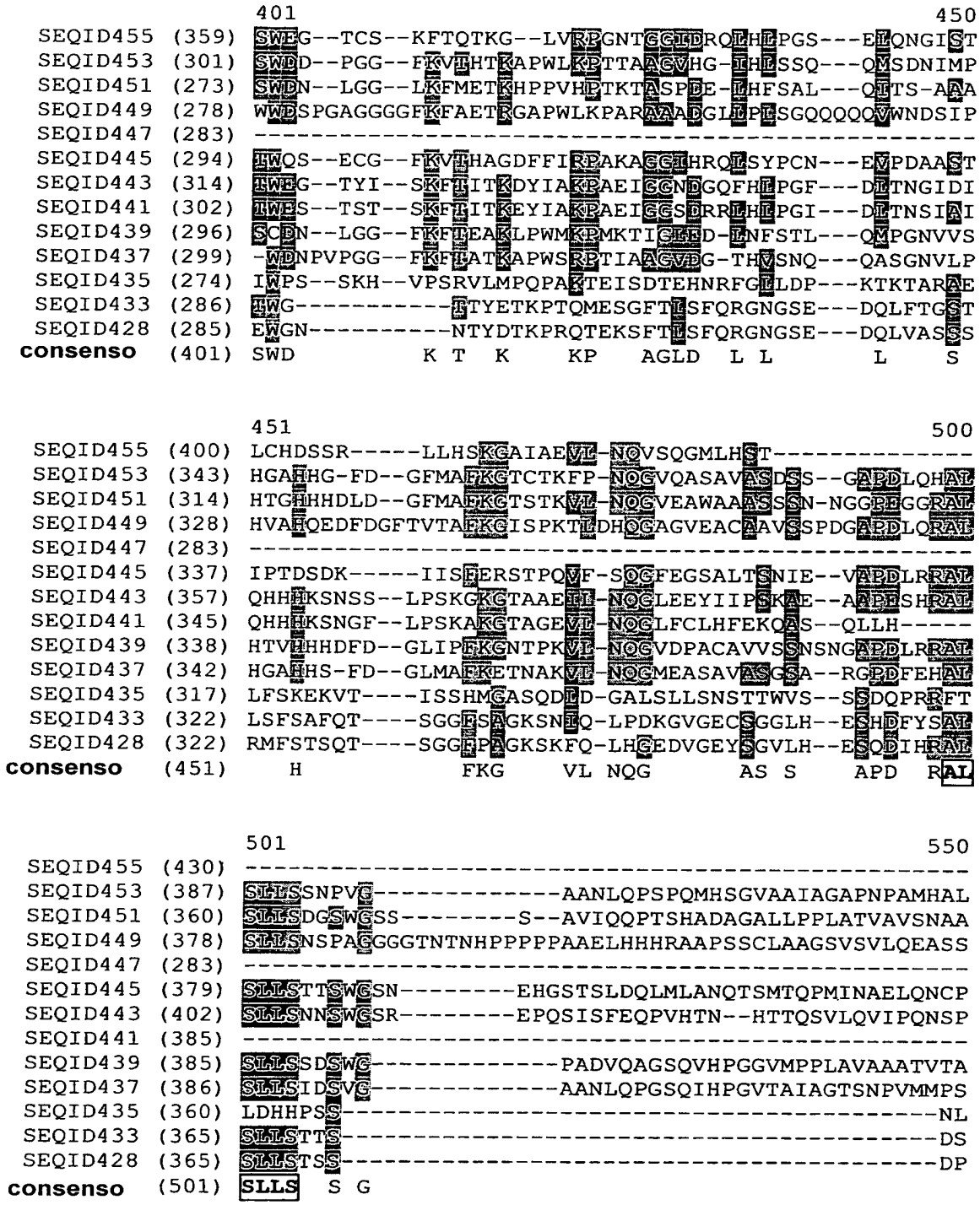


FIGURA 28 CONT

```
                    551                                                 600
  SEQID455  (430)  ---SEHWRTELPSTDSRVHILPSQGDGNTHFQEFKLFKAPYESGFSSGYS
  SEQID453  (424)  GSSTGLWLDGSQPLDDHPRFQVFERLG--DHDSELQLPKPSYDHASHFDR
  SEQID451  (400)  ----AAAGHPLDPSPGR--FWPQDDHPPLVDGPATQIPELAHLRIW----
  SEQID449  (428)  PGLWQDGGGSAALGHHAHARLQALDAAMATAQELLQLPRPPPSYGGSSSH
  SEQID447  (283)  --------------------------------------------------
  SEQID445  (421)  VASSENARMEQASLESRVHSLDLHDNGNVQLQEFQLLKAPYATGCFYSNQ
  SEQID443  (442)  LASSEYWRTEQPSTDSQVHTLTSH-------------------------
  SEQID441  (385)  --------------------------------------------------
  SEQID439  (422)  PTNPVSVMHALHPSTGGGGFWQDGDDPPPLD-HASQAQAFMHPGNGSSSG
  SEQID437  (423)  ---PAIWQGG-LSLDQQAQFQAFDRLGNDDDEDHLQLPKPSYDNS-HYDQ
  SEQID435  (369)  QPVAHRSAAQLNSVSGYWQPDPPAVEGPTALHRNGVGQFNENYFSLNQFY
  SEQID433  (374)  QGIKHTPVAEPPPIFGTFPSHFI-------------------------
  SEQID428  (374)  LAQPHVQPFSLLCSYDVVPK-----------------------------
consenso    (551)

                    601                 620
  SEQID455  (477)  EYRELRYPIPRSNARVMKQT
  SEQID453  (472)  MH------------------
  SEQID451  (440)  --------------------
  SEQID449  (478)  YDLMR---------------
  SEQID447  (283)  --------------------
  SEQID445  (471)  FS------------------
  SEQID443  (466)  --------------------
  SEQID441  (385)  --------------------
  SEQID439  (471)  YGHLH---------------
  SEQID437  (468)  MN------------------
  SEQID435  (419)  N-------------------
  SEQID433  (397)  --------------------
  SEQID428  (394)  --------------------
consenso    (601)
```

FIGURA 28 CONT

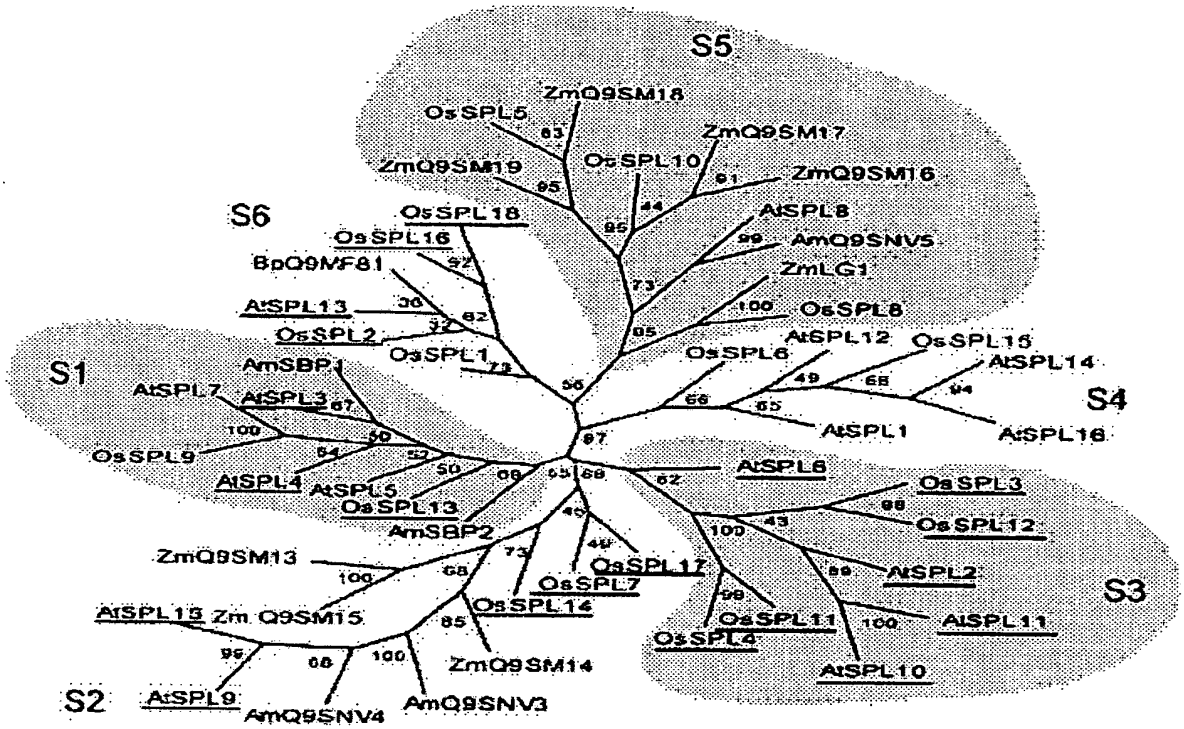


FIGURA 29

3'	aua	cga	gu g	aga	gaa	gac	agc	5'	*osmiR156l*
3'	aca	cga	ga g	aga	gaa	gac	agu	5'	*osmiR156k*
3'	ca	cga	gu g	aga	gaa	gac	agu	5'	*osmiR156(a -j)*
	• | |	| | |	| • |	| | |	| | |	| | |	| | |		
5'	cau	gcu	ccc	ucu	cuu	cug	uca	3'	*OsSPL13**
5'	gau	gcu	ccc	ucu	cuu	cug	uca	3'	*OsSPL3*
5'	ugu	gcu	cuc	ucu	cuu	cug	uca	3'	*OsSPL2*
5'	ugu	gcu	cuc	ucu	cuu	cug	uca	3'	*OsSPL14*
5'	ugu	gcu	cuc	ucu	cuu	cug	uca	3'	*OsSPL16*
5'	ugu	gcu	cuc	ucu	cuu	cug	uca	3'	*OsSPL17*
5'	ugu	gcu	cuc	ucu	cuu	cug	uca	3'	*OsSPL18*
5'	cgu	gcu	cuc	ucu	cuu	cug	uca	3'	*OsSPL4**
5'	cgu	gcu	cuc	ucu	cuu	cug	uca	3'	*OsSPL11*
5'	cgu	gcu	cuc	ucu	cuu	cug	uca	3'	*OsSPL12*
5'	ggu	gcu	cuc	ucu	cuu	cug	uca	3'	*OsSPL7*
A(G/C)		A	L	S	L	L	S		Aminoácido

FIGURA 20A

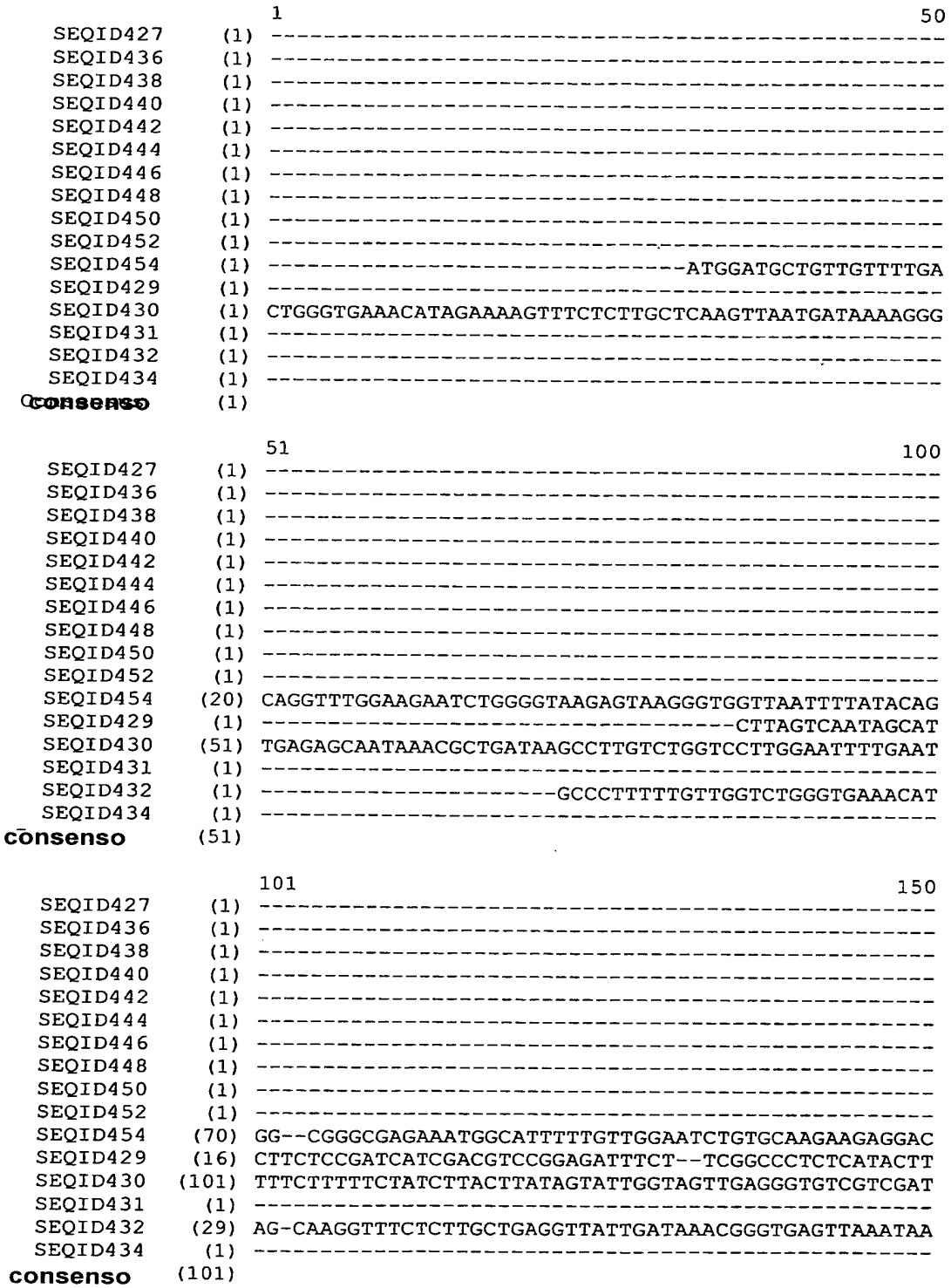

1 50
SEQID427 (1) ----------
SEQID436 (1) ----------
SEQID438 (1) ----------
SEQID440 (1) ----------
SEQID442 (1) ----------
SEQID444 (1) ----------
SEQID446 (1) ----------
SEQID448 (1) ----------
SEQID450 (1) ----------
SEQID452 (1) ----------
SEQID454 (1) ----------ATGGATGCTGTTGTTTTGA
SEQID429 (1) ----------
SEQID430 (1) CTGGGTGAAACATAGAAAAGTTTCTCTTGCTCAAGTTAATGATAAAAGGG
SEQID431 (1) ----------
SEQID432 (1) ----------
SEQID434 (1) ----------
consenso (1)
51 100
SEQID427 (1) ----------
SEQID436 (1) ----------
SEQID438 (1) ----------
SEQID440 (1) ----------
SEQID442 (1) ----------
SEQID444 (1) ----------
SEQID446 (1) ----------
SEQID448 (1) ----------
SEQID450 (1) ----------
SEQID452 (1) ----------
SEQID454 (20) CAGGTTTGGAAGAATCTGGGGTAAGAGTAAGGGTGGTTAATTTTATACAG
SEQID429 (1) ----------CTTAGTCAATAGCAT
SEQID430 (51) TGAGAGCAATAAACGCTGATAAGCCTTGTCTGGTCCTTGGAATTTTGAAT
SEQID431 (1) ----------
SEQID432 (1) ----------GCCCTTTTTGTTGGTCTGGGTGAAACAT
SEQID434 (1) ----------
consenso (51)
101 150
SEQID427 (1) ----------
SEQID436 (1) ----------
SEQID438 (1) ----------
SEQID440 (1) ----------
SEQID442 (1) ----------
SEQID444 (1) ----------
SEQID446 (1) ----------
SEQID448 (1) ----------
SEQID450 (1) ----------
SEQID452 (1) ----------
SEQID454 (70) GG--CGGGCGAGAAATGGCATTTTTGTTGGAATCTGTGCAAGAAGAGGAC
SEQID429 (16) CTTCTCCGATCATCGACGTCCGGAGATTTCT--TCGGCCCTCTCATACTT
SEQID430 (101) TTTCTTTTTCTATCTTACTTATAGTATTGGTAGTTGAGGGTGTCGTCGAT
SEQID431 (1) ----------
SEQID432 (29) AG-CAAGGTTTCTCTTGCTGAGGTTATTGATAAACGGGTGAGTTAAATAA
SEQID434 (1) ----------
consenso (101)

FIGURA 30B

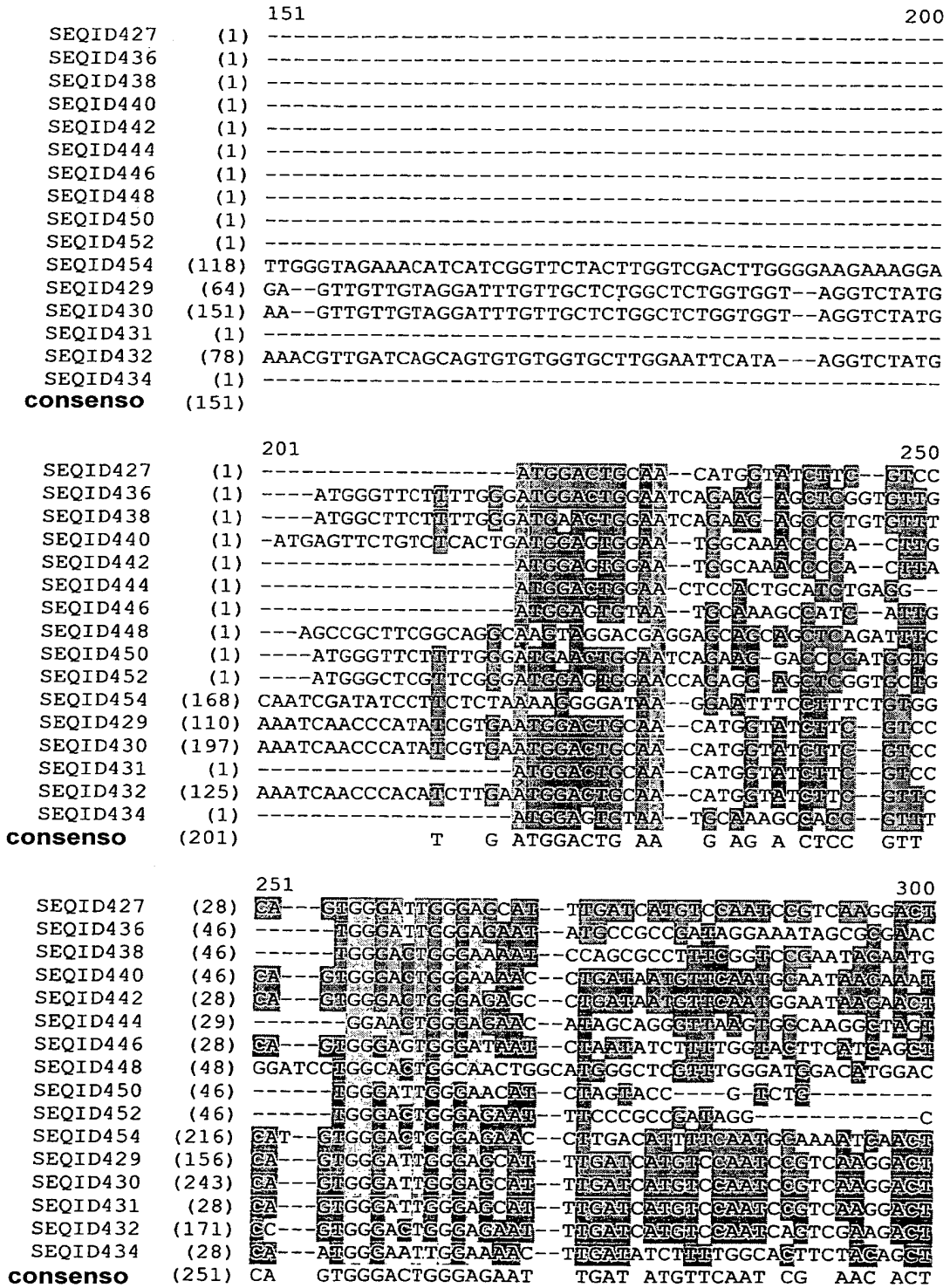


FIGURA 30B CONT.

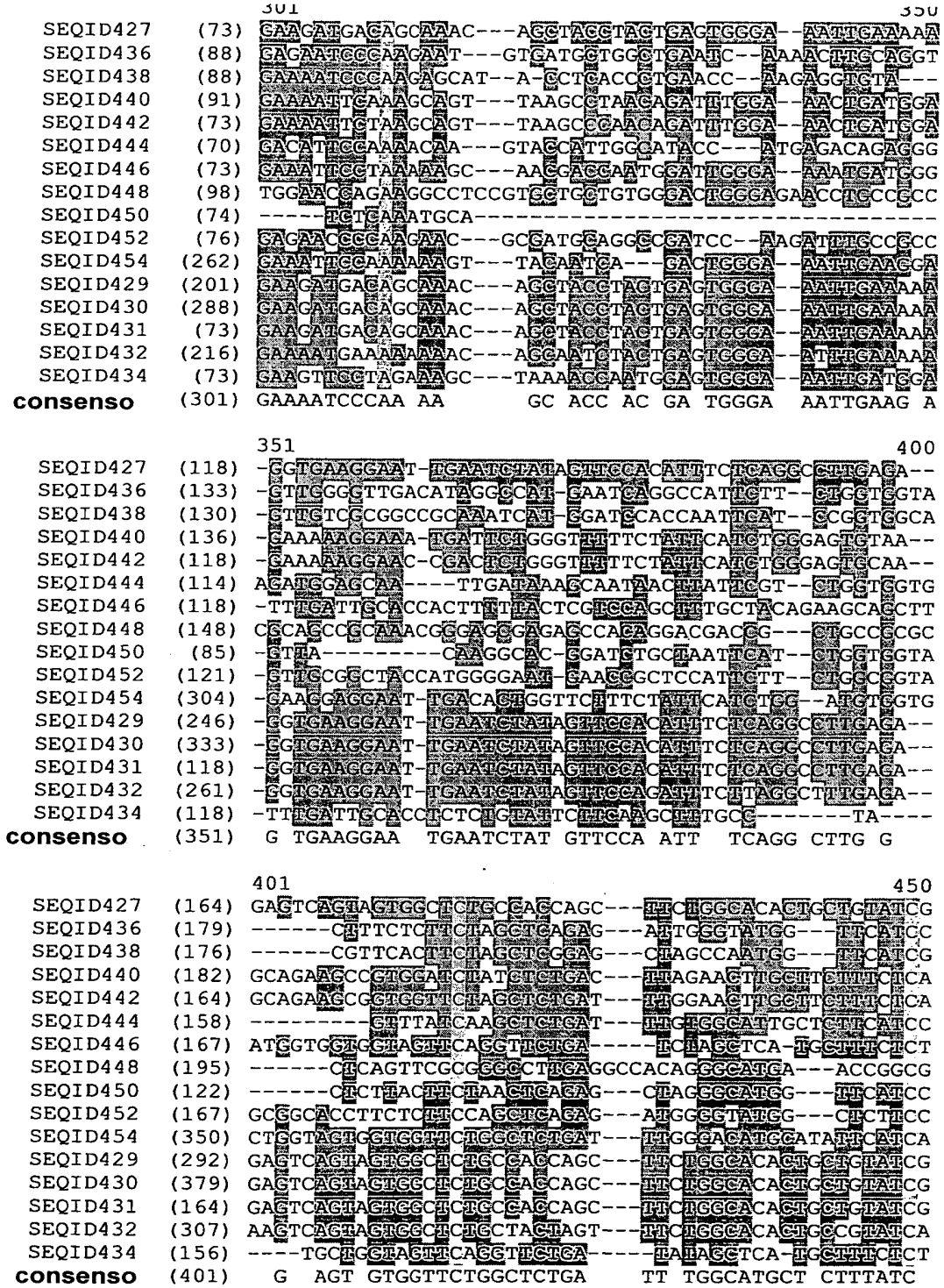


FIGURA 30B CONT.

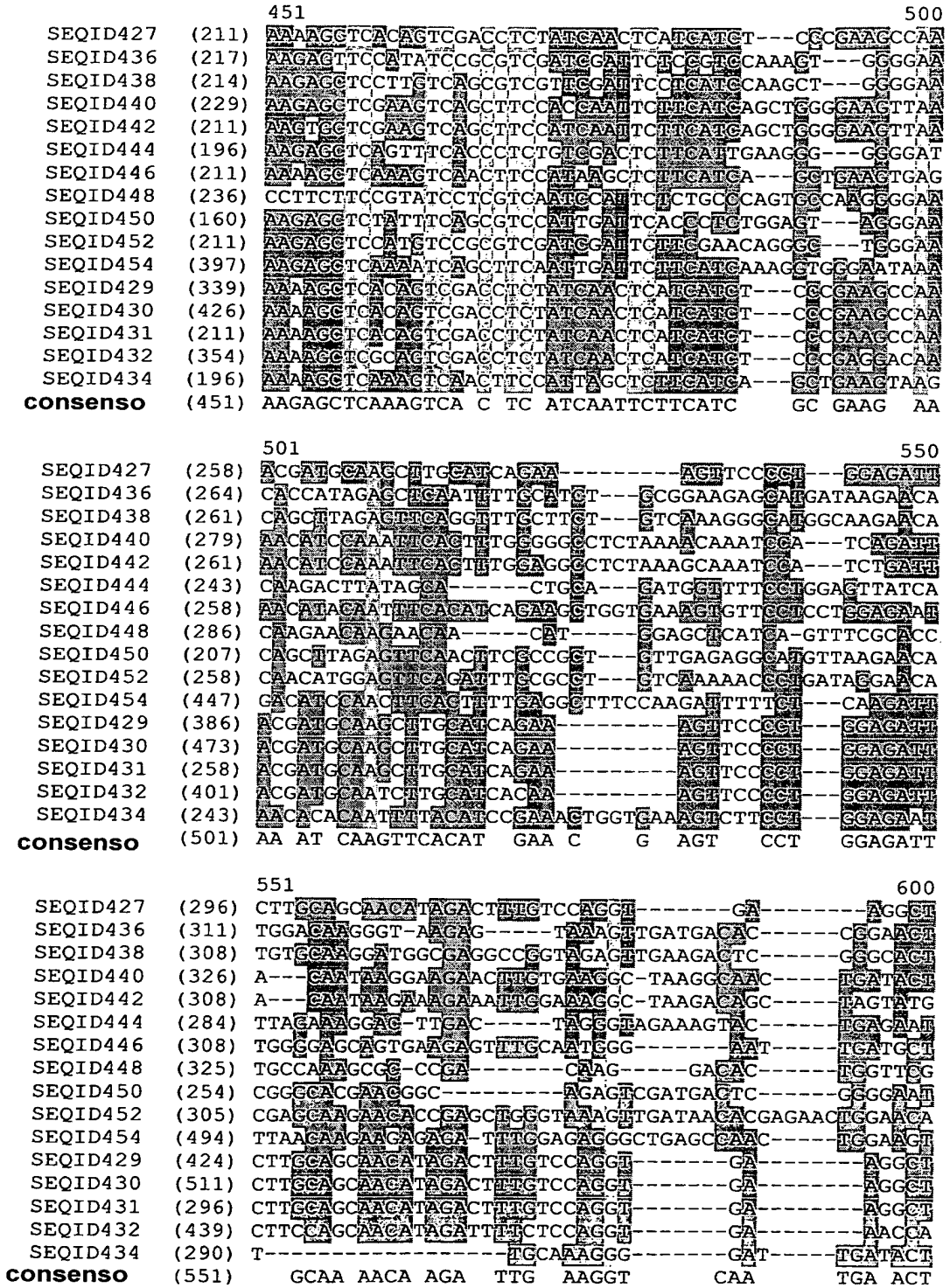


FIGURA 30B CONT.

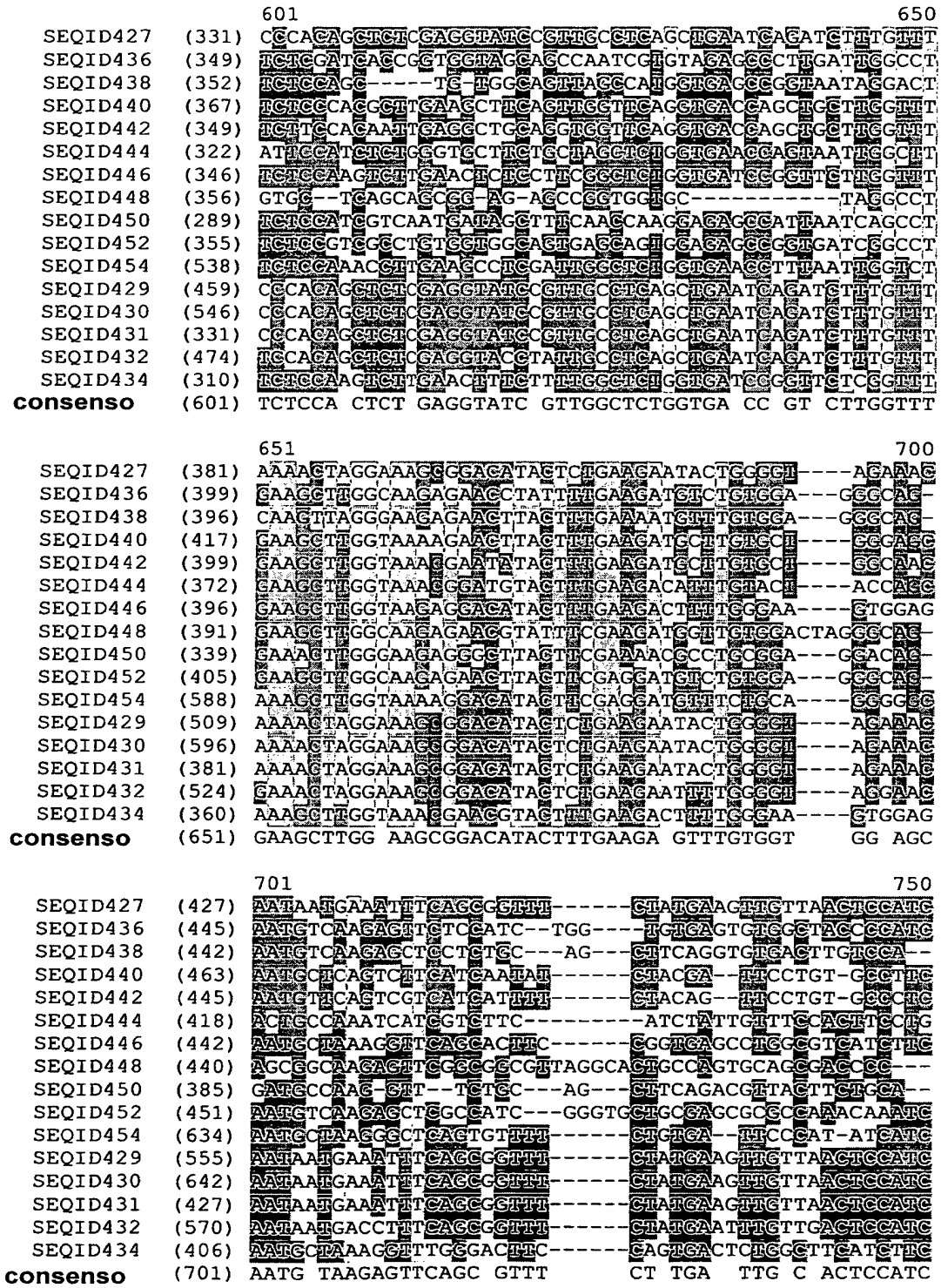


FIGURA 30B CONT.

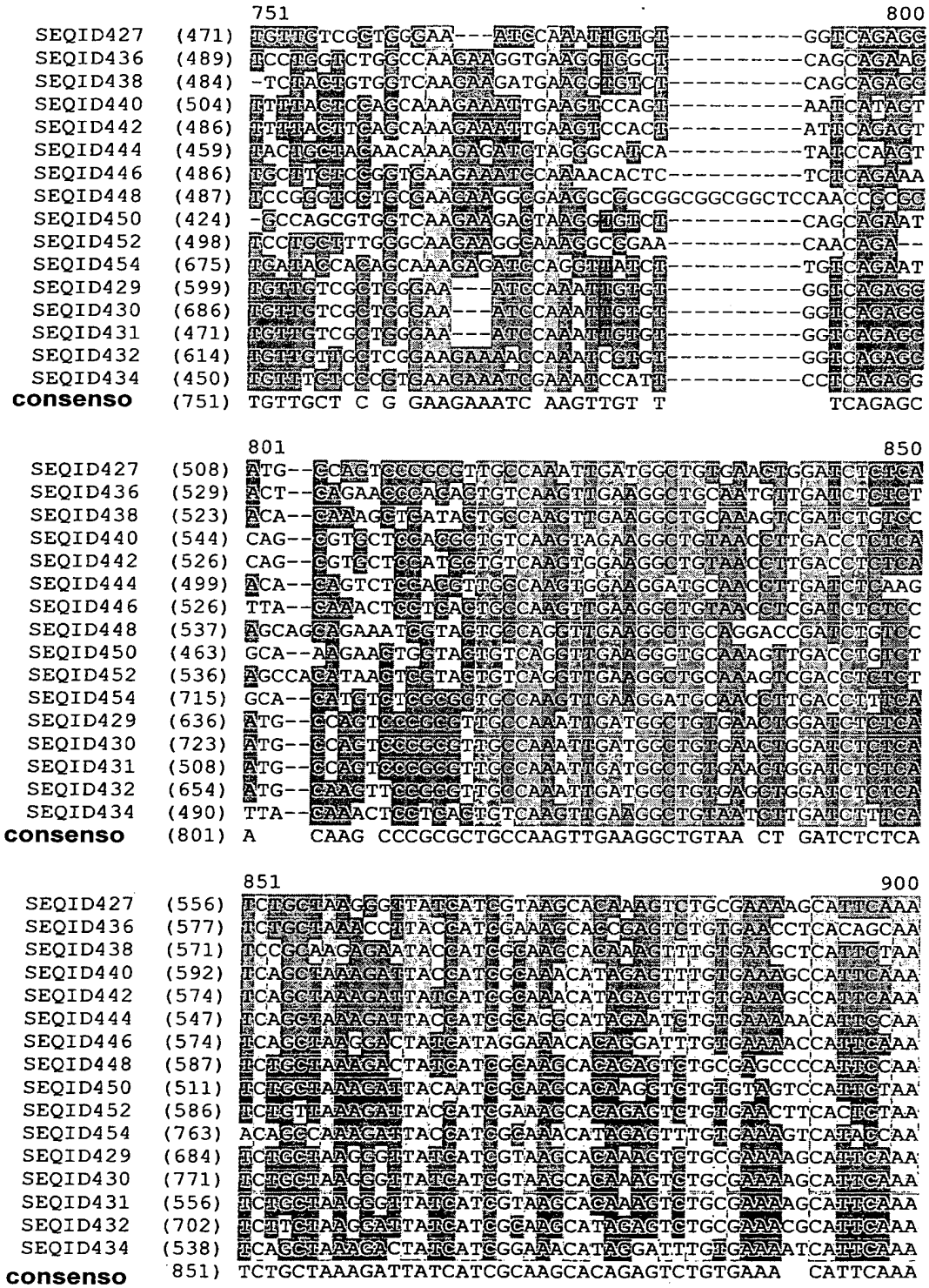


FIGURA 30B CONT.

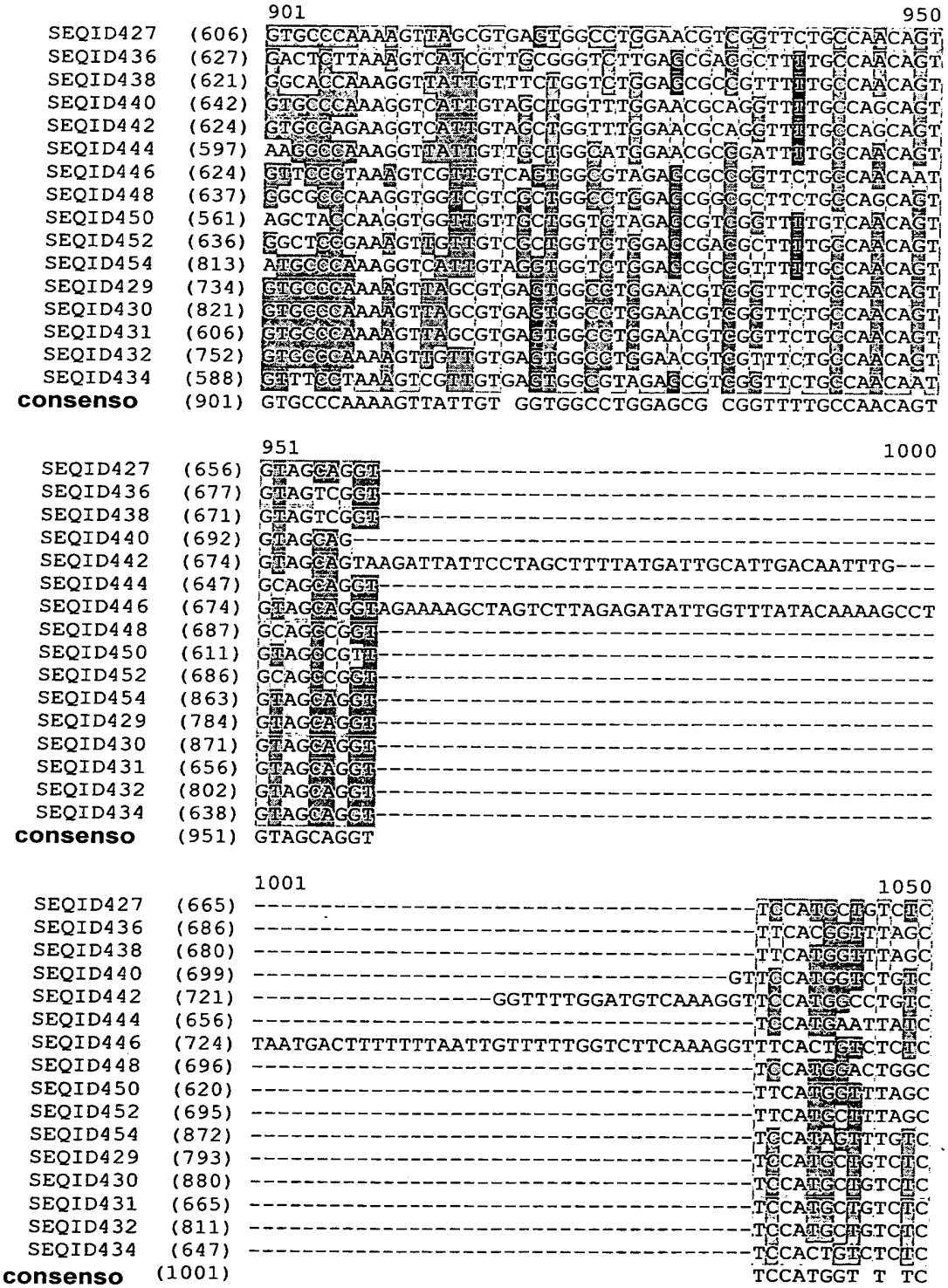


FIGURA 30B CONT.

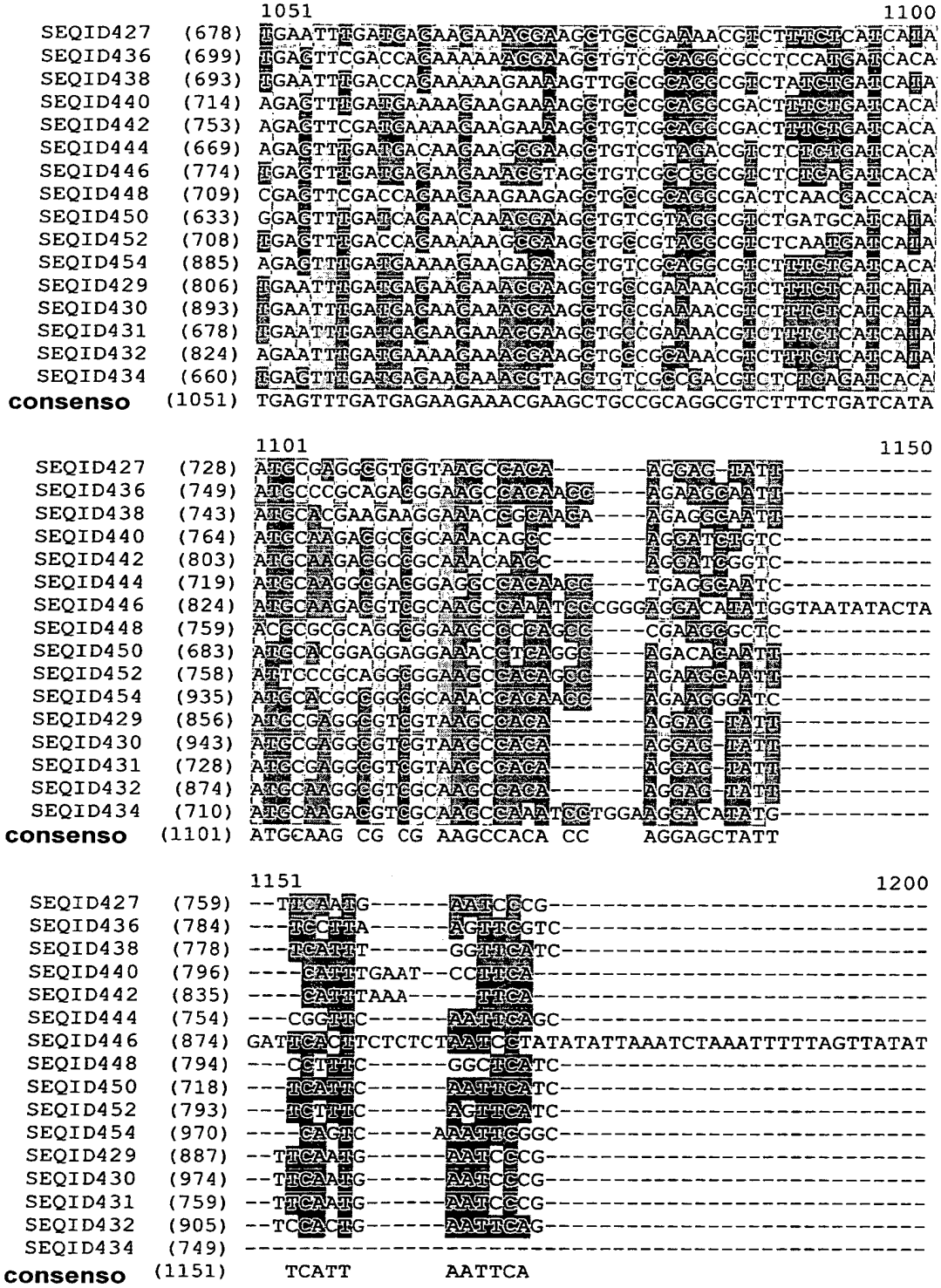


FIGURA 30B CONT.

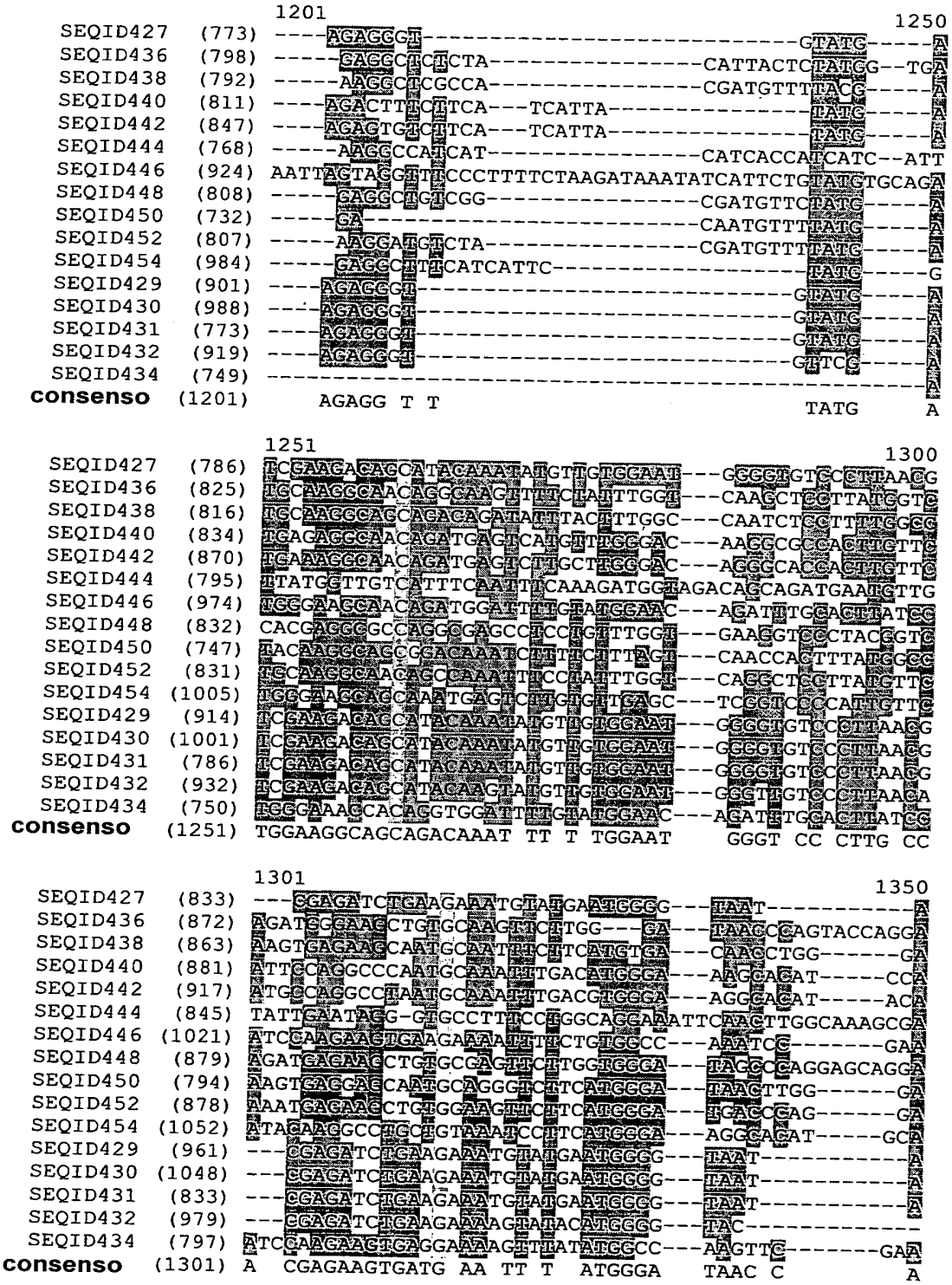


FIGURA 30B CONT.

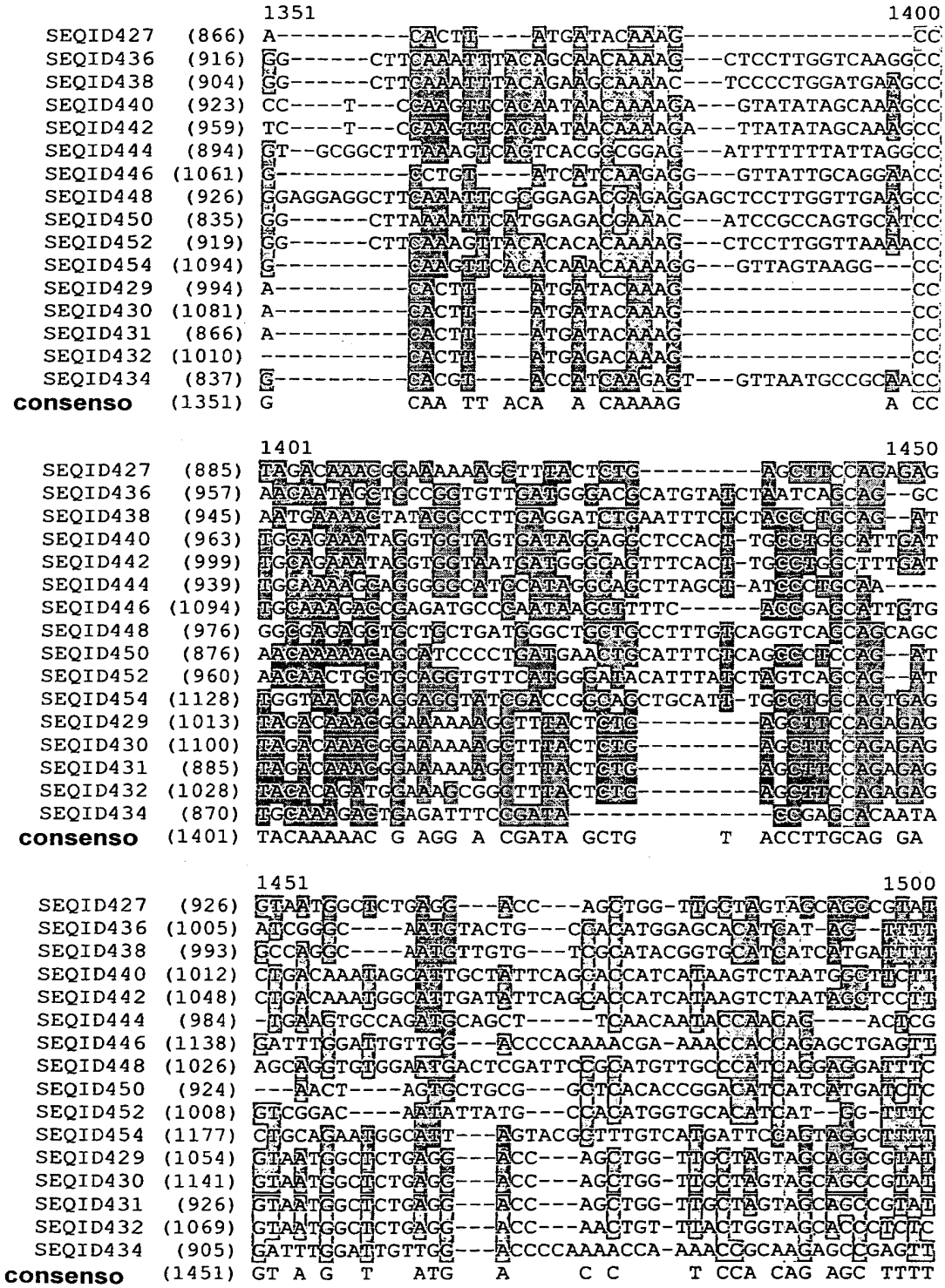


FIGURA 30B CONT.

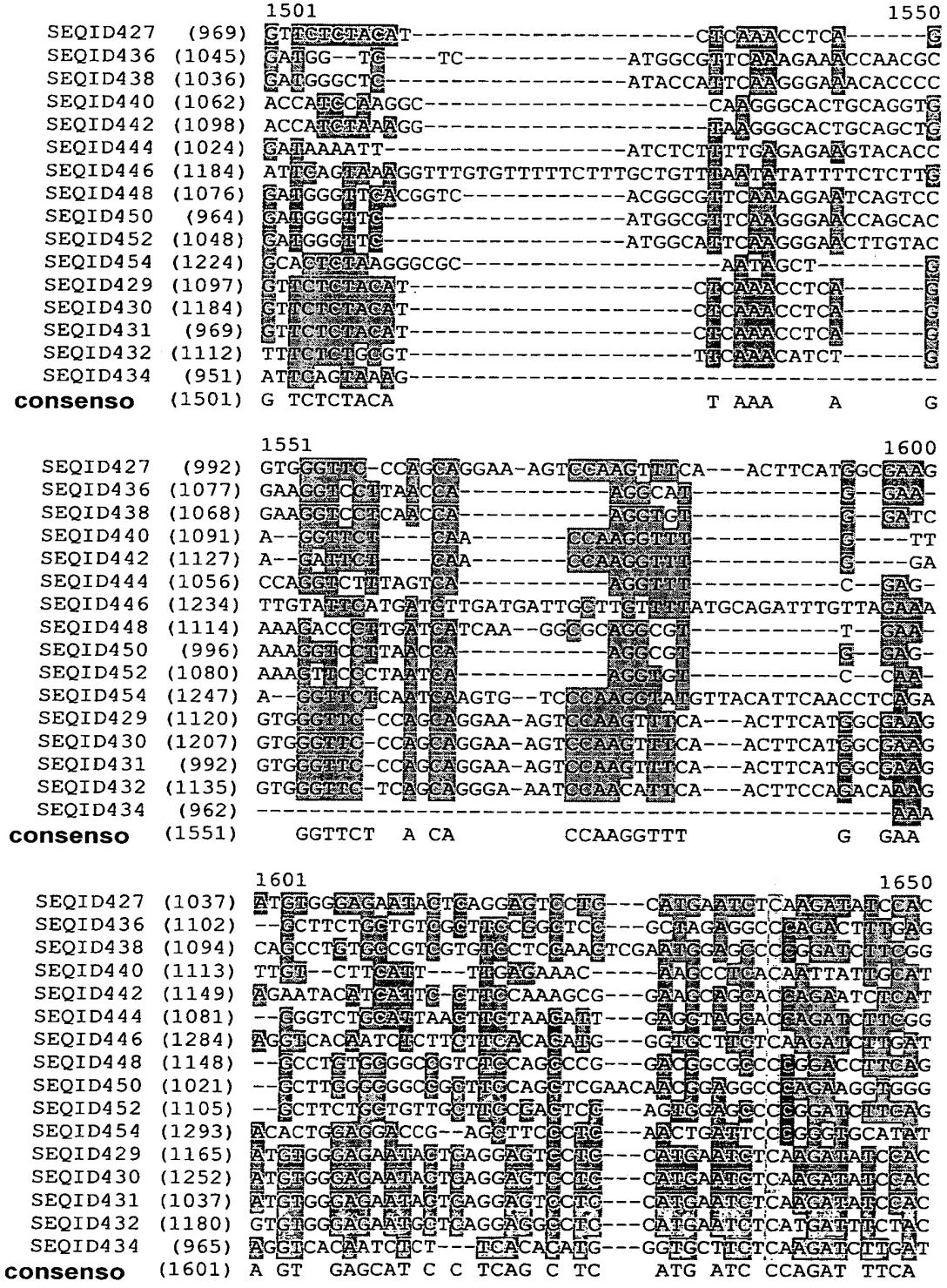


FIGURA 30B CONT.

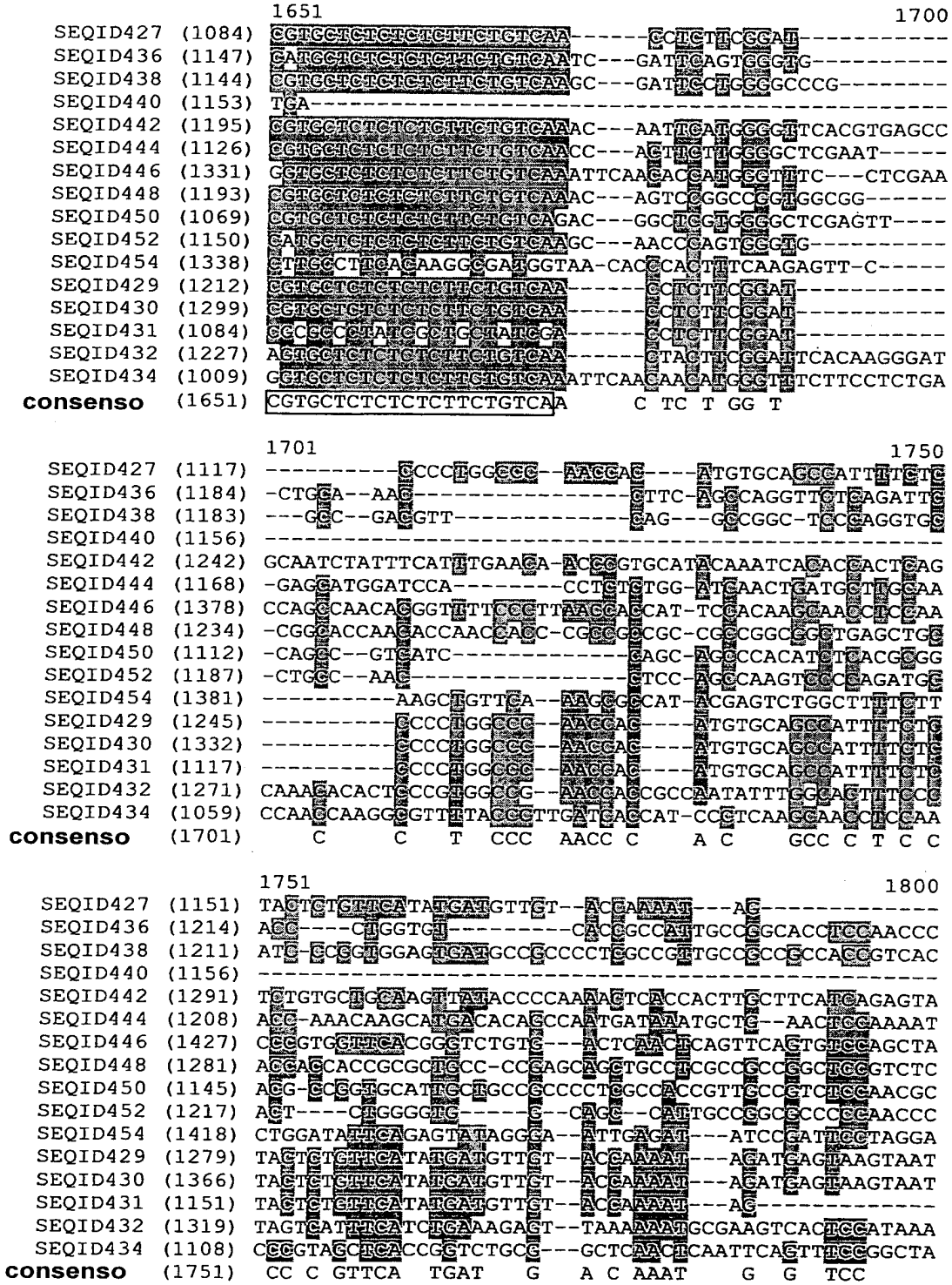


FIGURA 30B CONT.

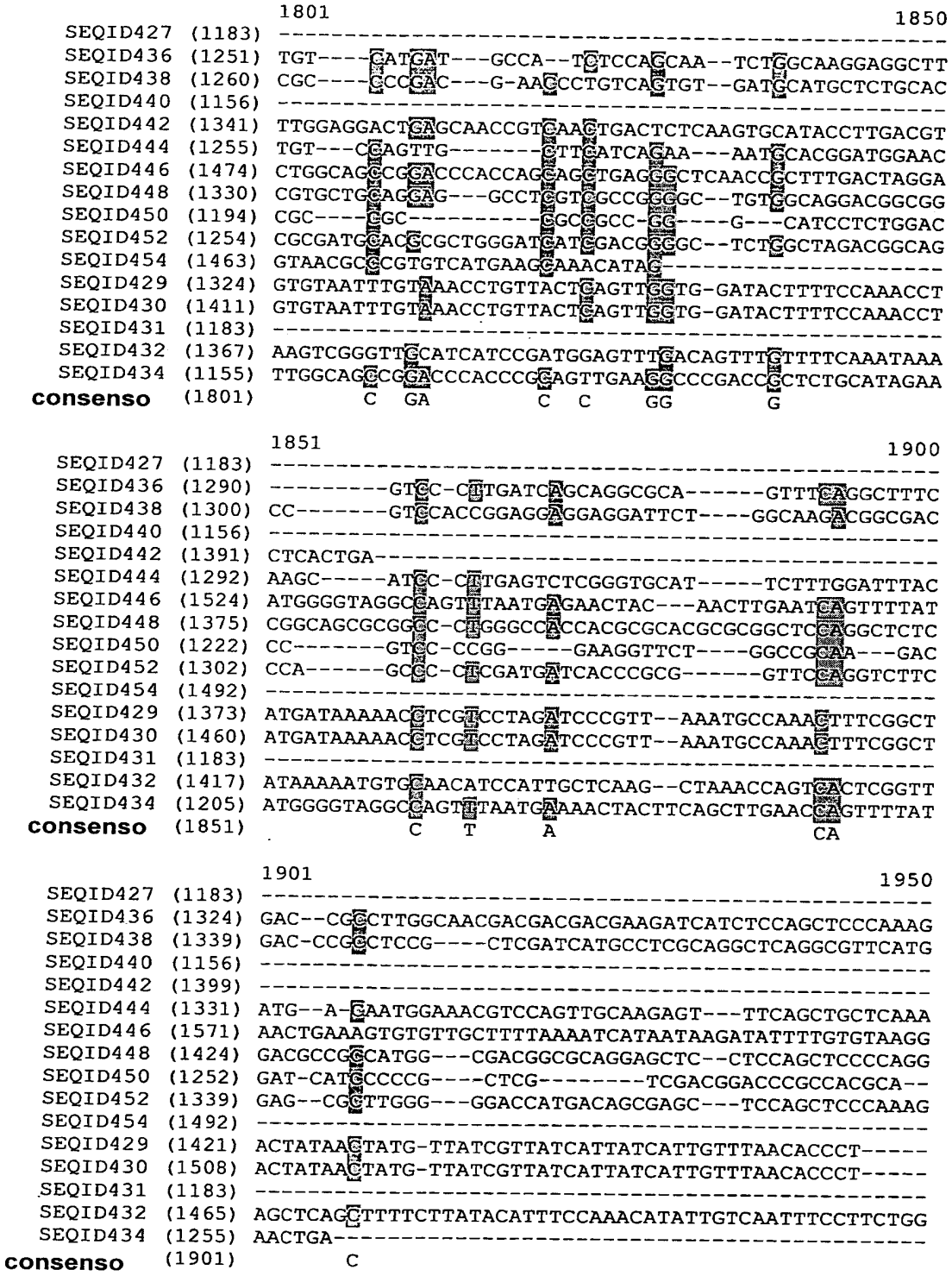


FIGURA 30B CONT.

```
                    1951                                                2000
SEQID427  (1183)  --------------------------------------------------
SEQID436  (1372)  CCT------TCCTATGACAACTCC------CACTATGATCAGATGAACTG
SEQID438  (1384)  CATCCTGGCAATGGCAGCAGCTCC--G----GCTATGGCCATCTGCACTG
SEQID440  (1156)  --------------------------------------------------
SEQID442  (1399)  --------------------------------------------------
SEQID444  (1375)  GCACCCTATGCTACCGGCTGCTTTT------ATTCCAATCAATTCAGCTA
SEQID446  (1621)  ATCAGGTGAGCCTAGTGGTAACGTTTAAATAGGAGCTGTGAAACTTGCTA
SEQID448  (1469)  CCGCCGCCATCGTACGGCGGCTCCTCGTCCCACTACGACCTGATGCGCTG
SEQID450  (1287)  GATTCCGGAGCTGGC-GCACCTCC--G----GATATGGTGA---------
SEQID452  (1381)  CCT------TCCTACGACCATGCCTCG---CACTTCGACCGGATGCACTG
SEQID454  (1492)  --------------------------------------------------
SEQID429  (1465)  --------------------------------------------------
SEQID430  (1552)  --------------------------------------------------
SEQID431  (1183)  --------------------------------------------------
SEQID432  (1515)  ATTCTGTTGGTATGACAACTTTGTTACTCTGTCAAACATGTCTACGACTA
SEQID434  (1261)  --------------------------------------------------
consenso  (1951)

                    2001                                                2050
SEQID427  (1183)  --------------------------------------------------
SEQID436  (1410)  A-------------------------------------------------
SEQID438  (1428)  A-------------------------------------------------
SEQID440  (1156)  --------------------------------------------------
SEQID442  (1399)  --------------------------------------------------
SEQID444  (1419)  A-------------------------------------------------
SEQID446  (1671)  AAGACATCAACTCCTCTCGTCTTTCTTTTGTCCATTGATTTTCTTGGGTT
SEQID448  (1519)  ATGCTTCCTTCCTTGTGGATTGTGGTGGAGAACTGGAGATGGAGCGGGAG
SEQID450  (1321)  --------------------------------------------------
SEQID452  (1422)  A-------------------------------------------------
SEQID454  (1492)  --------------------------------------------------
SEQID429  (1465)  --------------------------------------------------
SEQID430  (1552)  --------------------------------------------------
SEQID431  (1183)  --------------------------------------------------
SEQID432  (1565)  TTTTAAAACCATTTCAGAGATT--------------------------
SEQID434  (1261)  --------------------------------------------------
consenso  (2001)

                    2051                                                2100
SEQID427  (1183)  --------------------------------------------------
SEQID436  (1411)  --------------------------------------------------
SEQID438  (1429)  --------------------------------------------------
SEQID440  (1156)  --------------------------------------------------
SEQID442  (1399)  --------------------------------------------------
SEQID444  (1420)  --------------------------------------------------
SEQID446  (1721)  AGGGAGAGTTGTGACAGTTAAGTATTATAAATCATCTTGTCAAATTATTT
SEQID448  (1569)  CGCCGCCGTGCCCCCGGTTGGCGCGTGGAAATCAGTCAGACGTGGCTTCT
SEQID450  (1321)  --------------------------------------------------
SEQID452  (1423)  --------------------------------------------------
SEQID454  (1492)  --------------------------------------------------
SEQID429  (1465)  --------------------------------------------------
SEQID430  (1552)  --------------------------------------------------
SEQID431  (1183)  --------------------------------------------------
SEQID432  (1587)  --------------------------------------------------
SEQID434  (1261)  --------------------------------------------------
consenso  (2051)
```

FIGURA 30B CONT.

```
                    2101                                                2150
SEQID427  (1183)  --------------------------------------------------
SEQID436  (1411)  --------------------------------------------------
SEQID438  (1429)  --------------------------------------------------
SEQID440  (1156)  --------------------------------------------------
SEQID442  (1399)  --------------------------------------------------
SEQID444  (1420)  --------------------------------------------------
SEQID446  (1771)  ATACATTAGAATTTTCAGATATTGATGGTGTAGCAACTGAAAAATTGATC
SEQID448  (1619)  GTGGCGTCGTCTCTGCAACGCGGGAGCGGGATTTTGCTTGTGTTCTGAAC
SEQID450  (1321)  --------------------------------------------------
SEQID452  (1423)  --------------------------------------------------
SEQID454  (1492)  --------------------------------------------------
SEQID429  (1465)  --------------------------------------------------
SEQID430  (1552)  --------------------------------------------------
SEQID431  (1183)  --------------------------------------------------
SEQID432  (1587)  --------------------------------------------------
SEQID434  (1261)  --------------------------------------------------
consenso  (2101)

                    2151                                                2200
SEQID427  (1183)  --------------------------------------------------
SEQID436  (1411)  --------------------------------------------------
SEQID438  (1429)  --------------------------------------------------
SEQID440  (1156)  --------------------------------------------------
SEQID442  (1399)  --------------------------------------------------
SEQID444  (1420)  --------------------------------------------------
SEQID446  (1821)  AAACGCATCATAATATACGTATGATAAGGTTATTTGCATCTGAACTTTTT
SEQID448  (1669)  GAACAGAAAAAAAAAA----------------------------------
SEQID450  (1321)  --------------------------------------------------
SEQID452  (1423)  --------------------------------------------------
SEQID454  (1492)  --------------------------------------------------
SEQID429  (1465)  --------------------------------------------------
SEQID430  (1552)  --------------------------------------------------
SEQID431  (1183)  --------------------------------------------------
SEQID432  (1587)  --------------------------------------------------
SEQID434  (1261)  --------------------------------------------------
consenso  (2151)

                    2201                                                2250
SEQID427  (1183)  --------------------------------------------------
SEQID436  (1411)  --------------------------------------------------
SEQID438  (1429)  --------------------------------------------------
SEQID440  (1156)  --------------------------------------------------
SEQID442  (1399)  --------------------------------------------------
SEQID444  (1420)  --------------------------------------------------
SEQID446  (1871)  TTATATTCGCATTAGTCAGCATTTTGTGTTATACTAGATTTGGATCCGTG
SEQID448  (1685)  --------------------------------------------------
SEQID450  (1321)  --------------------------------------------------
SEQID452  (1423)  --------------------------------------------------
SEQID454  (1492)  --------------------------------------------------
SEQID429  (1465)  --------------------------------------------------
SEQID430  (1552)  --------------------------------------------------
SEQID431  (1183)  --------------------------------------------------
SEQID432  (1587)  --------------------------------------------------
SEQID434  (1261)  --------------------------------------------------
consenso  (2201)
```

FIGURA 30B CONT.

```
                    2251                                                2300
SEQID427  (1183)  ----------------------------------------------------------
SEQID436  (1411)  ----------------------------------------------------------
SEQID438  (1429)  ----------------------------------------------------------
SEQID440  (1156)  ----------------------------------------------------------
SEQID442  (1399)  ----------------------------------------------------------
SEQID444  (1420)  ----------------------------------------------------------
SEQID446  (1921)  CGTTGCAACAGGTTTTATTTGATATTACCATCCATTGATTTTTTTTTTTT
SEQID448  (1685)  ----------------------------------------------------------
SEQID450  (1321)  ----------------------------------------------------------
SEQID452  (1423)  ----------------------------------------------------------
SEQID454  (1492)  ----------------------------------------------------------
SEQID429  (1465)  ----------------------------------------------------------
SEQID430  (1552)  ----------------------------------------------------------
SEQID431  (1183)  ----------------------------------------------------------
SEQID432  (1587)  ----------------------------------------------------------
SEQID434  (1261)  ----------------------------------------------------------
consenso  (2251)

                    2301                                                2350
SEQID427  (1183)  ----------------------------------------------------------
SEQID436  (1411)  ----------------------------------------------------------
SEQID438  (1429)  ----------------------------------------------------------
SEQID440  (1156)  ----------------------------------------------------------
SEQID442  (1399)  ----------------------------------------------------------
SEQID444  (1420)  ----------------------------------------------------------
SEQID446  (1971)  TTTGCAAACAAATAATTTGGTAATTCCATTAACCGCCACAGTCGTTTTCT
SEQID448  (1685)  ----------------------------------------------------------
SEQID450  (1321)  ----------------------------------------------------------
SEQID452  (1423)  ----------------------------------------------------------
SEQID454  (1492)  ----------------------------------------------------------
SEQID429  (1465)  ----------------------------------------------------------
SEQID430  (1552)  ----------------------------------------------------------
SEQID431  (1183)  ----------------------------------------------------------
SEQID432  (1587)  ----------------------------------------------------------
SEQID434  (1261)  ----------------------------------------------------------
consenso  (2301)

                    2351              2372
SEQID427  (1183)  ----------------------
SEQID436  (1411)  ----------------------
SEQID438  (1429)  ----------------------
SEQID440  (1156)  ----------------------
SEQID442  (1399)  ----------------------
SEQID444  (1420)  ----------------------
SEQID446  (2021)  GAGCAAAACCATGTTAAAATCT
SEQID448  (1685)  ----------------------
SEQID450  (1321)  ----------------------
SEQID452  (1423)  ----------------------
SEQID454  (1492)  ----------------------
SEQID429  (1465)  ----------------------
SEQID430  (1552)  ----------------------
SEQID431  (1183)  ----------------------
SEQID432  (1587)  ----------------------
SEQID434  (1261)  ----------------------
consenso  (2351)
```

FIGURA 30B CONT.

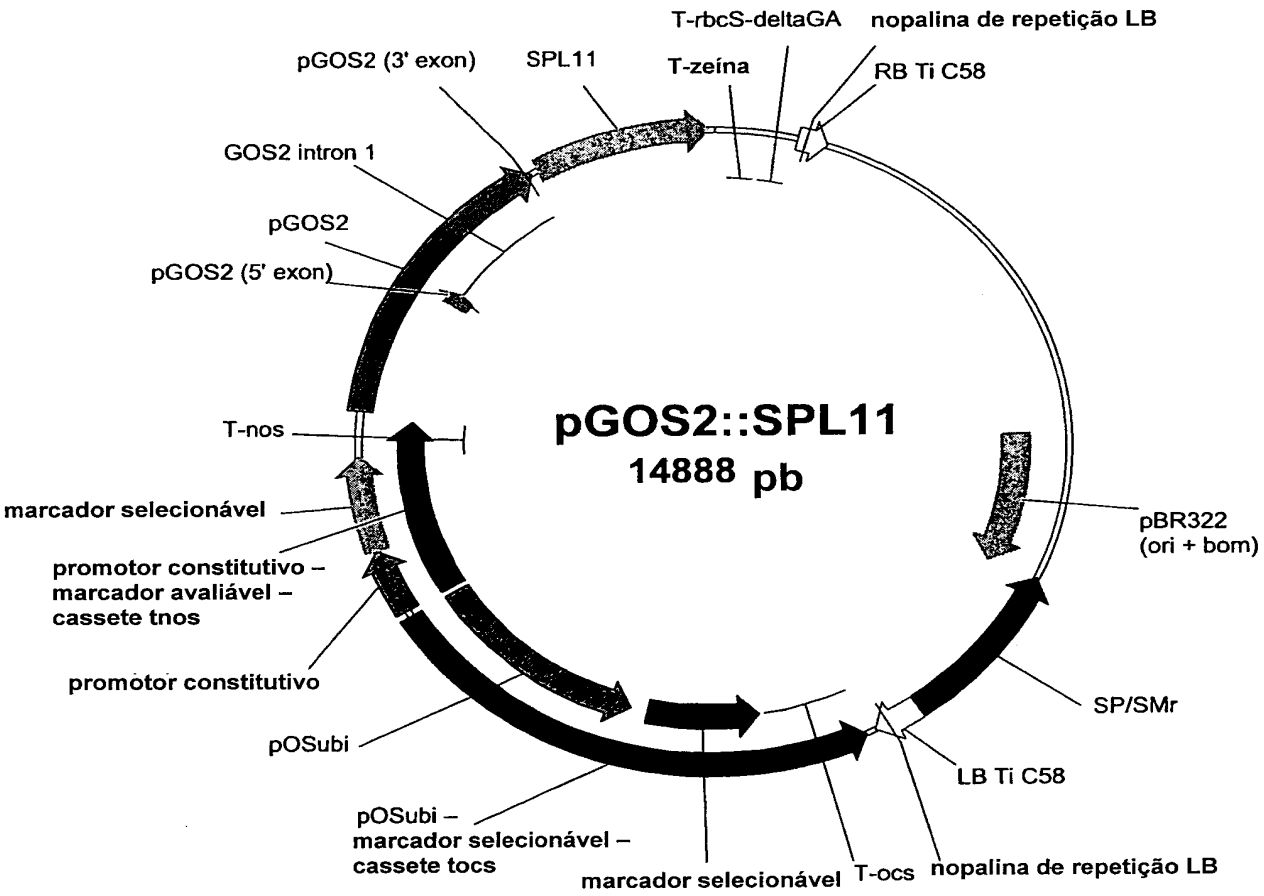

T-rbcS-deltaGA
nopalina de repetição LB
pGOS2 (3' exon)
SPL11
T-zeína
RB Ti C58
GOS2 intron 1
pGOS2
pGOS2 (5' exon)
pGOS2::SPL11
14888 pb
T-nos
marcador selecionável
promotor constitutivo –
marcador avaliável –
cassete tnos
promotor constitutivo
pOSubi
pOSubi –
marcador selecionável –
cassete tocs
marcador selecionável
T-ocs
nopalina de repetição LB
LB Ti C58
SP/SMr
pBR322
(ori + bom)

FIGURA 31A

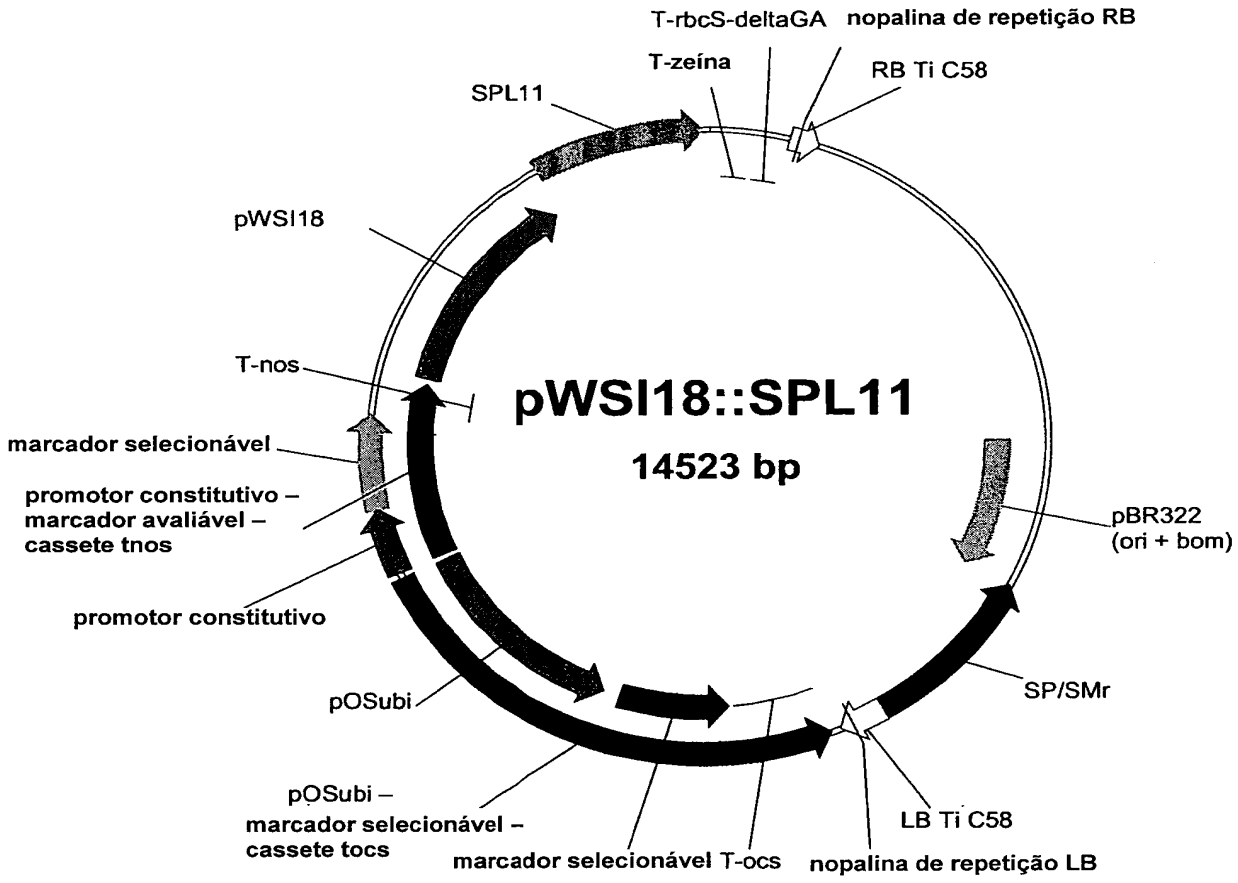

T-rbcS-deltaGA
nopalina de repetição RB
T-zeína
RB Ti C58
SPL11
pWSI18
T-nos
pWSI18::SPL11
14523 bp
marcador selecionável
promotor constitutivo –
marcador avaliável –
cassete tnos
promotor constitutivo
pBR322
(ori + bom)
pOSubi
SP/SMr
pOSubi –
marcador selecionável –
cassete tocs
marcador selecionável T-ocs
LB Ti C58
nopalina de repetição LB

FIGURA 31B

SEQ ID NO: 427, DNA - *Arabidopsis thaliana*

ATGGACTGCAACATGGTATCTTCGTCCCAGTGGGATTGGGAGCATTTGATCATGTCCAATCCGTCA
AGGACTGAAGATGACAGCAAACAGCTACCTACTGAGTGGGAAATTGAAAAAGGTGAAGGAATTGAA
TCTATAGTTCCACATTTCTCAGGCCTTGAGAGAGTCAGTAGTGGCTCTGCCACCAGCTTCTGGCAC
ACTGCTGTATCGAAAAGCTCACAGTCGACCTCTATCAACTCATCATCTCCCGAAGCCAAACGATGC
AAGCTTGCATCAGAAAGTTCCCCTGGAGATTCTTGCAGCAACATAGACTTTGTCCAGGTGAAGGCT
CCCACAGCTCTCGAGGTATCCGTTGCCTCAGCTGAATCAGATCTTTGTTTAAAACTAGGAAAGCGG
ACATACTCTGAAGAATACTGGGGTAGAAACAATAATGAAATTTCAGCGGTTTCTATGAAGTTGTTA
ACTCCATCTGTTGTCGCTGGGAAATCCAAATTGTGTGGTCAGAGCATGCCAGTCCCGCGTTGCCAA
ATTGATGGCTGTGAACTGGATCTCTCATCTGCTAAGGGTTATCATCGTAAGCACAAAGTCTGCGAA
AAGCATTCAAAGTGCCCAAAAGTTAGCGTGAGTGGCCTGGAACGTCGGTTCTGCCAACAGTGTAGC
AGGTTCCATGCTGTCTCTGAATTTGATGAGAAGAAACGAAGCTGCCGAAAACGTCTTTCTCATCAT
AATGCGAGGCGTCGTAAGCCACAAGGAGTATTTTCAATGAATCCCGAGAGGGTGTATGATCGAAGA
CAGCATACAAATATGTTGTGGAATGGGGTGTCCCTTAACGCGAGATCTGAAGAAATGTATGAATGG
GGTAATAACACTTATGATACAAAGCCTAGACAAACGGAAAAAAGCTTTACTCTGAGCTTCCAGAGA
GGTAATGGCTCTGAGGACCAGCTGGTTGCTAGTAGCAGCCGTATGTTCTCTACATCTCAAACCTCA
GGTGGGTTCCCAGCAGGAAAGTCCAAGTTTCAACTTCATGGCGAAGATGTGGGAGAATACTCAGGA
GTCCTCCATGAATCTCAAGATATCCACCGTGCTCTCTCTCTTCTGTCAACCTCTTCGGATCCCCTG
GCCCAACCACATGTGCAGCCATTTTCTCTACTCTGTTCATATGATGTTGTACCAAAATAG

SEQ ID NO: 428, proteína - *Arabidopsis thaliana*

MDCNMVSSSQWDWEHLIMSNPSRTEDDSKQLPTEWEIEKGEGIESIVPHFSGLERVSSGSATSFWH
TAVSKSSQSTSINSSSPEAKRCKLASESSPGDSCSNIDFVQVKAPTALEVSVASAESDLCLKLGKR
TYSEEYWGRNNNEISAVSMKLLTPSVVAGKSKLCGQSMPVPRCQIDGCELDLSSAKGYHRKHKVCE
KHSKCPKVSVSGLERRFCQQCSRFHAVSEFDEKKRSCRKRLSHHNARRRKPQGVFSMNPERVYDRR
QHTNMLWNGVSLNARSEEMYEWGNNTYDTKPRQTEKSFTLSFQRGNGSEDQLVASSSRMFSTSQTS
GGFPAGKSKFQLHGEDVGEYSGVLHESQDIHRALSLLSTSSDPLAQPHVQPFSLLCSYDVVPK

SEQ ID NO: 429, DNA - *Arabidopsis thaliana*

CTTAGTCAATAGCATCTTCTCCGATCATCGACGTCCGGAGATTTCTTCGGCCCTCTCATACTTGAG
TTGTTGTAGGATTTGTTGCTCTGGCTCTGGTGGTAGGTCTATGAAATCAACCCATATCGTGAATGG
ACTGCAACATGGTATCTTCGTCCCAGTGGGATTGGGAGCATTTGATCATGTCCAATCCGTCAAGGA
CTGAAGATGACAGCAAACAGCTACCTACTGAGTGGGAAATTGAAAAAGGTGAAGGAATTGAATCTA
TAGTTCCACATTTCTCAGGCCTTGAGAGAGTCAGTAGTGGCTCTGCCACCAGCTTCTGGCACACTG
CTGTATCGAAAAGCTCACAGTCGACCTCTATCAACTCATCATCTCCCGAAGCCAAACGATGCAAGC
TTGCATCAGAAAGTTCCCCTGGAGATTCTTGCAGCAACATAGACTTTGTCCAGGTGAAGGCTCCCA
CAGCTCTCGAGGTATCCGTTGCCTCAGCTGAATCAGATCTTTGTTTAAAACTAGGAAAGCGGACAT
ACTCTGAAGAATACTGGGGTAGAAACAATAATGAAATTTCAGCGGTTTCTATGAAGTTGTTAACTC
CATCTGTTGTCGCTGGGAAATCCAAATTGTGTGGTCAGAGCATGCCAGTCCCGCGTTGCCAAATTG
ATGGCTGTGAACTGGATCTCTCATCTGCTAAGGGTTATCATCGTAAGCACAAAGTCTGCGAAAAGC
ATTCAAAGTGCCCAAAAGTTAGCGTGAGTGGCCTGGAACGTCGGTTCTGCCAACAGTGTAGCAGGT
TCCATGCTGTCTCTGAATTTGATGAGAAGAAACGAAGCTGCCGAAAACGTCTTTCTCATCATAATG
CGAGGCGTCGTAAGCCACAAGGAGTATTTTCAATGAATCCCGAGAGGGTGTATGATCGAAGACAGC
ATACAAATATGTTGTGGAATGGGGTGTCCCTTAACGCGAGATCTGAAGAAATGTATGAATGGGGTA
ATAACACTTATGATACAAAGCCTAGACAAACGGAAAAAAGCTTTACTCTGAGCTTCCAGAGAGGTA
ATGGCTCTGAGGACCAGCTGGTTGCTAGTAGCAGCCGTATGTTCTCTACATCTCAAACCTCAGGTG
GGTTCCCAGCAGGAAAGTCCAAGTTTCAACTTCATGGCGAAGATGTGGGAGAATACTCAGGAGTCC
TCCATGAATCTCAAGATATCCACCGTGCTCTCTCTCTTCTGTCAACCTCTTCGGATCCCCTGGCCC

FIGURA 32

AACCACATGTGCAGCCATTTTCTCTACTCTGTTCATATGATGTTGTACCAAAATAGATGAGTAAGT
AATGTGTAATTTGTAAACCTGTTACTCAGTTGGTGGATACTTTTCCAAACCTATGATAAAAACCTC
GTCCTAGATCCCGTTAAATGCCAAACTTTCGGCTACTATAACTATGTTATCGTTATCATTATCATT
GTTTAACACCCT

SEQ ID NO: 430, DNA - *Arabidopsis thaliana*

CTGGGTGAAACATAGAAAAGTTTCTCTTGCTCAAGTTAATGATAAAAGGGTGAGAGCAATAAACGC
TGATAAGCCTTGTCTGGTCCTTGGAATTTTGAATTTTCTTTTTCTATCTTACTTATAGTATTGGTA
GTTGAGGGTGTCGTCGATAAGTTGTTGTAGGATTTGTTGCTCTGGCTCTGGTGGTAGGTCTATGAA
ATCAACCCATATCGTGAATGGACTGCAACATGGTATCTTCGTCCCAGTGGGATTGGGAGCATTTGA
TCATGTCCAATCCGTCAAGGACTGAAGATGACAGCAAACAGCTACCTACTGAGTGGGAAATTGAAA
AAGGTGAAGGAATTGAATCTATAGTTCCACATTTCTCAGGCCTTGAGAGAGTCAGTAGTGGCTCTG
CCACCAGCTTCTGGCACACTGCTGTATCGAAAAGCTCACAGTCGACCTCTATCAACTCATCATCTC
CCGAAGCCAAACGATGCAAGCTTGCATCAGAAAGTTCCCCTGGAGATTCTTGCAGCAACATAGACT
TTGTCCAGGTGAAGGCTCCCACAGCTCTCGAGGTATCCGTTGCCTCAGCTGAATCAGATCTTTGTT
TAAAACTAGGAAAGCGGACATACTCTGAAGAATACTGGGGTAGAAACAATAATGAAATTTCAGCGG
TTTCTATGAAGTTGTTAACTCCATCTGTTGTCGCTGGGAAATCCAAATTGTGTGGTCAGAGCATGC
CAGTCCCGCGTTGCCAAATTGATGGCTGTGAACTGGATCTCTCATCTGCTAAGGGTTATCATCGTA
AGCACAAAGTCTGCGAAAAGCATTCAAAGTGCCCAAAAGTTAGCGTGAGTGGCCTGGAACGTCGGT
TCTGCCAACAGTGTAGCAGGTTCCATGCTGTCTCTGAATTTGATGAGAAGAAACGAAGCTGCCGAA
AACGTCTTTCTCATCATAATGCGAGGCGTCGTAAGCCACAAGGAGTATTTTCAATGAATCCCGAGA
GGGTGTATGATCGAAGACAGCATACAAATATGTTGTGGAATGGGGTGTCCCTTAACGCGAGATCTG
AAGAAATGTATGAATGGGGTAATAACACTTATGATACAAAGCCTAGACAAACGGAAAAAAGCTTTA
CTCTGAGCTTCCAGAGAGGTAATGGCTCTGAGGACCAGCTGGTTGCTAGTAGCAGCCGTATGTTCT
CTACATCTCAAACCTCAGGTGGGTTCCCAGCAGGAAAGTCCAAGTTTCAACTTCATGGCGAAGATG
TGGGAGAATACTCAGGAGTCCTCCATGAATCTCAAGATATCCACCGTGCTCTCTCTCTTCTGTCAA
CCTCTTCGGATCCCCTGGCCCAACCACATGTGCAGCCATTTTCTCTACTCTGTTCATATGATGTTG
TACCAAAATAGATGAGTAAGTAATGTGTAATTTGTAAACCTGTTACTCAGTTGGTGGATACTTTTC
CAAACCTATGATAAAAACCTCGTCCTAGATCCCGTTAAATGCCAAACTTTCGGCTACTATAACTAT
GTTATCGTTATCATTATCATTGTTTAACACCCT

SEQ ID NO: 431, DNA - *Arabidopsis thaliana*

ATGGACTGCAACATGGTATCTTCGTCCCAGTGGGATTGGGAGCATTTGATCATGTCCAATCCGTCA
AGGACTGAAGATGACAGCAAACAGCTACCTACTGAGTGGGAAATTGAAAAAGGTGAAGGAATTGAA
TCTATAGTTCCACATTTCTCAGGCCTTGAGAGAGTCAGTAGTGGCTCTGCCACCAGCTTCTGGCAC
ACTGCTGTATCGAAAAGCTCACAGTCGACCTCTATCAACTCATCATCTCCCGAAGCCAAACGATGC
AAGCTTGCATCAGAAAGTTCCCCTGGAGATTCTTGCAGCAACATAGACTTTGTCCAGGTGAAGGCT
CCCACAGCTCTCGAGGTATCCGTTGCCTCAGCTGAATCAGATCTTTGTTTAAAACTAGGAAAGCGG
ACATACTCTGAAGAATACTGGGGTAGAAACAATAATGAAATTTCAGCGGTTTCTATGAAGTTGTTA
ACTCCATCTGTTGTCGCTGGGAAATCCAAATTGTGTGGTCAGAGCATGCCAGTCCCGCGTTGCCAA
ATTGATGGCTGTGAACTGGATCTCTCATCTGCTAAGGGTTATCATCGTAAGCACAAAGTCTGCGAA
AAGCATTCAAAGTGCCCAAAAGTTAGCGTGAGTGGCCTGGAACGTCGGTTCTGCCAACAGTGTAGC
AGGTTCCATGCTGTCTCTGAATTTGATGAGAAGAAACGAAGCTGCCGAAAACGTCTTTCTCATCAT
AATGCGAGGCGTCGTAAGCCACAAGGAGTATTTTCAATGAATCCCGAGAGGGTGTATGATCGAAGA
CAGCATACAAATATGTTGTGGAATGGGGTGTCCCTTAACGCGAGATCTGAAGAAATGTATGAATGG
GGTAATAACACTTATGATACAAAGCCTAGACAAACGGAAAAAAGCTTTACTCTGAGCTTCCAGAGA
GGTAATGGCTCTGAGGACCAGCTGGTTGCTAGTAGCAGCCGTATGTTCTCTACATCTCAAACCTCA
GGTGGGTTCCCAGCAGGAAAGTCCAAGTTTCAACTTCATGGCGAAGATGTGGGAGAATACTCAGGA
GTCCTCCATGAATCTCAAGATATCCACCGCGCCCTATCGCTGCTATCGACCTCTTCGGATCCCCTG
GCCCAACCACATGTGCAGCCATTTTCTCTACTCTGTTCATATGATGTTGTACCAAAATAG

FIGURA 32 CONT.

SEQ ID NO: 432, DNA - *Arabidopsis thaliana*

GCCCTTTTTGTTGGTCTGGGTGAAACATAGCAAGGTTTCTCTTGCTGAGGTTATTGATAAACGGGT
GAGTTAAATAAAAACGTTGATCAGCAGTGTGTGGTGCTTGGAATTCATAAGGTCTATGAAATCAAC
CCACATCTTGAATGGACTGCAACATGGTATCTTCGTTCCCGTGGGACTGGGAGAATTTGATCATGT
CCAATCAGTCGAAGACTGAAAATGAAAAAAAACAGCAATCTACTGAGTGGGAATTTGAAAAAGGTG
AAGGAATTGAATCTATAGTTCCAGATTTCTTAGGCTTTGAGAAAGTCAGTAGTGGCTCTGCTACTA
GTTTCTGGCACACTGCCGTATCAAAAAGCTCGCAGTCGACCTCTATCAACTCATCATCTCCCGAGG
ACAAACGATGCAATCTTGCATCACAAAGTTCCCCTGGAGATTCTTCCAGCAACATAGATTTTCTCC
AGGTGAAACCATCCACAGCTCTCGAGGTACCTATTGCCTCAGCTGAATCAGATCTTTGTTTGAAAC
TAGGAAAGCGGACATACTCTGAAGAATTTTGGGGTAGGAACAATAATGACCTTTCAGCGGTTTCTA
TGAATTTGTTGACTCCATCTGTTGTTGCTCGGAAGAAAACCAAATCGTGTGGTCAGAGCATGCAAG
TTCCGCGTTGCCAAATTGATGGCTGTGAGCTGGATCTCTCATCTTCTAAGGATTATCATCGCAAGC
ATAGAGTCTGCGAAACGCATTCAAAGTGCCCAAAAGTTGTTGTGAGTGGCCTGGAACGTCGTTTCT
GCCAACAGTGTAGCAGGTTCCATGCTGTCTCAGAATTTGATGAAAAGAAACGAAGCTGCCGCAAAC
GTCTTTCTCATCATAATGCAAGGCGTCGCAAGCCACAAGGAGTATTTCCACTGAATTCAGAGAGGG
TGTTCGATCGAAGACAGCATACAAGTATGTTGTGGAATGGGTTGTCCCTTAACACGAGATCTGAAG
AAAAGTATACATGGGGTACCACTTATGAGACAAAGCCTACACAGATGGAAAGCGGCTTTACTCTGA
GCTTCCAGAGAGGTAATGGCTCTGAGGACCAACTGTTTACTGGTAGCACCCTCTCTTTCTCTGCGT
TTCAAACATCTGGTGGGTTCTCAGCAGGGAAATCCAACATTCAACTTCCAGACAAAGGTGTGGGAG
AATGCTCAGGAGGCCTCCATGAATCTCATGATTTCTACAGTGCTCTCTCTCTTCTGTCAACTACTT
CGGATTCACAAGGGATCAAACACACTCCCGTGGCCGAACCACCGCCAATATTTGGCACTTTCCCTA
GTCATTTCATCTGAAAGAGTTAAAAAATGCGAAGTCACTCCATAAAAAGTCGGGTTGCATCATCCG
ATGGAGTTTGACAGTTTGTTTTCAAATAAAATAAAAATGTGCAACATCCATTGCTCAAGCTAAACC
AGTCACTCGGTTAGCTCAGCTTTTCTTATACATTTCCAAACATATTGTCAATTTCCTTCTGGATTC
TGTTGGTATGACAACTTTGTTACTCTGTCAAACATGTCTACGACTATTTTAAAACCATTTCAGAGA
TT

SEQ ID NO: 433, proteína - *Arabidopsis thaliana*

MDCNMVSSFPWDWENLIMSNQSKTENEKKQQSTEWEFEKGEGIESIVPDFLGFEKVSSGSATSFWH
TAVSKSSQSTSINSSSPEDKRCNLASQSSPGDSSSNIDFLQVKPSTALEVPIASAESDLCLKLGKR
TYSEEFWGRNNNDLSAVSMNLLTPSVVARKKTKSCGQSMQVPRCQIDGCELDLSSSKDYHRKHRVC
ETHSKCPKVVVSGLERRFCQQCSRFHAVSEFDEKKRSCRKRLSHHNARRRKPQGVFPLNSERVFDR
RQHTSMLWNGLSLNTRSEEKYTWGTTYETKPTQMESGFTLSFQRGNGSEDQLFTGSTLSFSAFQTS
GGFSAGKSNIQLPDKGVGECSGGLHESHDFYSALSLLSTTSDSQGIKHTPVAEPPPIFGTFPSHFI

SEQ ID NO: 434, DNA - *Arabidopsis thaliana*

ATGGAGTGTAATGCAAAGCCACCGTTTCAATGGGAATTGGAAAACTTGATATCTTTTGGCACTTCT
ACAGCTGAAGTTCCTAGAAAGCTAAAACCAATGGAGTGGGAAATTGATGGATTTGATTGCACCTCT
CTGTATTCTTCAAGCTTTGCCTATGCTGGTAGTTCAGGTTCTGATATAGCTCATGCTTTCTCTAAA
AGCTCAAAGTCAACTTCCATTAGCTCTTCATCAGCTGAAGTAAGAACACACAATTTTACATCCGAA
ACTGGTGAAAGTCTTCCTGGAGAATTTGCAAAGGGGATTGATACTTCTCCAAGTCTTGAACTTTCT
TTTGGCTCTGGTGATCCGGTTCTCGGTTTAAAGCTTGGTAAACGAACGTACTTTGAAGACTTTTGG
GAAGTGGAGAATGCTAAAGGTTTGGGACTTCCAGTGACTCTGGCTTCATCTTCTGTTTCTCCCGTG
AAGAAATCGAAATCCATTCCTCAGAGGTTACAAACTCCTCACTGTCAAGTTGAAGGCTGTAATCTT
GATCTTTCATCAGCTAAAGACTATCATCGGAAACATAGGATTTGTGAAAATCATTCAAAGTTTCCT
AAAGTCGTTGTGAGTGGCGTAGAGCGTCGGTTCTGCCAACAATGTAGCAGGTTCCACTGTCTCTCT
GAGTTTGATGAGAAGAAACGTAGCTGTCGCCGACGTCTCTCAGATCACAATGCAAGACGTCGCAAG
CCAAATCCTGGAAGGACATATGATGGGAAACCACAGGTGGATTTTGTATGGAACAGATTTGCACTT

FIGURA 32 CONT.

ATCCATCCAAGAAGTGAGGAAAAGTTTATATGGCCAAGTTCGAAGCACGTACCATCAAGAGTGTTA
ATGCCGCAACCTGCAAAGACTGAGATTTCCGATACCGAGCACAATAGATTTGGATTGTTGGACCCC
AAAACCAAAACCGCAAGAGCCGAGTTATTCAGTAAAGAAAAGGTCACAATCTCTTCACACATGGGT
GCTTCTCAAGATCTTGATGGTGCTCTCTCTCTTCTGTCAAATTCAACAACATGGGTTTCTTCCTCT
GACCAACCAAGGCGTTTTACCCTTGATCACCATCCCTCAAGCAACCTCCAACCCGTAGCTCACCGG
TCTGCGGCTCAACTCAATTCAGTTTCCGGCTATTGGCAGCCGGACCCACCCGCAGTTGAAGGCCCG
ACCGCTCTGCATAGAAATGGGGTAGGCCAGTTTAATGAAAACTACTTCAGCTTGAACCAGTTTTAT
AACTGA

SEQ ID NO: 435, proteína - *Arabidopsis thaliana*
MECNAKPPFQWELENLISFGTSTAEVPRKLKPMEWEIDGFDCTSLYSSSFAYAGSSGSDIAHAFSK
SSKSTSISSSSAEVRTHNFTSETGESLPGEFAKGIDTSPSLELSFGSGDPVLGLKLGKRTYFEDFW
EVENAKGLGLPVTLASSSVSPVKKSKSIPQRLQTPHCQVEGCNLDLSSAKDYHRKHRICENHSKFP
KVVVSGVERRFCQQCSRFHCLSEFDEKKRSCRRRLSDHNARRRKPNPGRTYDGKPQVDFVWNRFAL
IHPRSEEKFIWPSSKHVPSRVLMPQPAKTEISDTEHNRFGLLDPKTKTARAELFSKEKVTISSHMG
ASQDLDGALSLLSNSTTWVSSSDQPRRFTLDHHPSSNLQPVAHRSAAQLNSVSGYWQPDPPAVEGP
TALHRNGVGQFNENYFSLNQFYN

SEQ ID NO: 436, DNA - *Oryza sativa*
ATGGGTTCTTTTGGGATGGACTGGAATCAGAAGAGCTCGGTGTTGTGGGATTGGGAGAATATGCCG
CCGATAGGAAATAGCGCGAACGAGAATCCCAAGAATGTGATGCTGGCTGAATCAAAACTTGCAGGT
GTTGGGGTTGACATAGGCCATGAATCAGGCCATTCTTCTGGTGGTACTTTCTCTTCTAGCTCAGAG
ATTGGGTATGGTTCATCCAAGAGTTCCATATCCGCGTCGATCGATTCTCCGTCCAAAGTGGGGAAC
ACCATAGAGCTCAATTTTGCATCTGCGGAAGAGCATGATAAGAACATGGACAAGGGTAAGAGTAAA
GTTGATGACACCGGAACTTCTCGATCACCGGTGGTAGCAGCCAATCGTGTAGAGCCCTTGATTGGC
CTGAAGCTTGGCAAGAGAACCTATTTTGAAGATGTCTGTGGAGGGCAGAATGTCAAGAGTTCTCCA
TCTGGTGTGAGTGTGGCTACCCCATCTCCTGGTCTGGCCAAGAAGGTGAAGGTGGCTCAGCAGAAC
ACTCAGAACCCACACTGTCAAGTTGAAGGCTGCAATGTTGATCTCTCTTCTGCTAAACCTTACCAT
CGAAAGCACCGAGTCTGTGAACCTCACAGCAAGACTCTTAAAGTCATCGTTGCGGGTCTTGAGCGA
CGCTTTTGCCAACAGTGTAGTCGGTTTCACGGTTTAGCTGAGTTCGACCAGAAAAAACGAAGCTGT
CGCAGGCGCCTCCATGATCACAATGCCCGCAGACGGAAGCCACAACCAGAAGCAATTTCCTTAAGT
TCGTCGAGGCTCTCTACATTACTCTATGGTGATGCAAGGCAACAGGCAAGTTTTCTATTTGGTCAA
GCTCCTTATGGTCAGATGGGAAGCTGTGCAAGTTCTTGGGATAACCCAGTACCAGGAGGCTTCAAA
TTTACAGCAACAAAAGCTCCTTGGTCAAGGCCAACAATAGCTGCCGGTGTTGATGGGACGCATGTA
TCTAATCAGCAGGCATCGGGCAATGTACTGCCACATGGAGCACATCATAGTTTTGATGGTCTCATG
GCGTTCAAAGAAACCAACGCGAAGGTCCTTAACCAAGGCATGGAAGCTTCTGCTGTCGCTTCCGGC
TCCGCTAGAGGCCCAGACTTTGAGCATGCTCTCTCTCTTCTGTCAATCGATTCAGTGGGTGCTGCA
AACCTTCAGCCAGGTTCTCAGATTCACCCTGGTGTCACCGCCATTGCCGGCACCTCCAACCCTGTC
ATGATGCCATCTCCAGCAATCTGGCAAGGAGGCTTGTCCCTTGATCAGCAGGCGCAGTTTCAGGCT
TTCGACCGCCTTGGCAACGACGACGACGAAGATCATCTCCAGCTCCCAAAGCCTTCCTATGACAAC
TCCCACTATGATCAGATGAACTGA

SEQ ID NO: 437, proteína - *Oryza sativa*
MGSFGMDWNQKSSVLWDWENMPPIGNSANENPKNVMLAESKLAGVGVDIGHESGHSSGGTFSSSSE
IGYGSSKSSISASIDSPSKVGNTIELNFASAEEHDKNMDKGKSKVDDTGTSRSPVVAANRVEPLIG
LKLGKRTYFEDVCGGQNVKSSPSGVSVATPSPGLAKKVKVAQQNTQNPHCQVEGCNVDLSSAKPYH
RKHRVCEPHSKTLKVIVAGLERRFCQQCSRFHGLAEFDQKKRSCRRRLHDHNARRRKPQPEAISLS

FIGURA 32 CONT.

SSRLSTLLYGDARQQASFLFGQAPYGQMGSCASSWDNPVPGGFKFTATKAPWSRPTIAAGVDGTHV
SNQQASGNVLPHGAHHSFDGLMAFKETNAKVLNQGMEASAVASGSARGPDFEHALSLLSIDSVGAA
NLQPGSQIHPGVTAIAGTSNPVMMPSPAIWQGGLSLDQQAQFQAFDRLGNDDDEDHLQLPKPSYDN
SHYDQMN

SEQ ID NO: 438, DNA - *Oryza sativa*

ATGGCTTCTTTTGGGATGAACTGGAATCAGAAGAGCCCTGTGTTTTGGGACTGGGAAAATCCAGCG
CCTTTCGGTCCGAATACAATGGAAAATCCCAAGAGCATACCTCACCCTGAACCAAGAGGTGTAGTT
GTCGCGGCCGCAAATCATGGATCCACCAATTCATCCGGTGGCACGTTCACTTCTAGCTCGGAGCTA
GCCAATGGTTCATCGAAGAGCTCCTTGTCAGCGTCGTTCGATTCCTCATCCAAGCTGGGGAACAGC
TTAGAGTTCAGGTTTGCTTCTGTCAAAGGGCATGGCAAGAACATGTGCAAGGATGGCGAGGCCGGT
AGAGTTGAAGACTCGGGCACTTCTCCAGCTGTGGCAGTTAGCCATGGTGAGCCGGTAATAGGACTC
AAGTTAGGGAAGAGAACTTACTTTGAAAATGTTTGTGGAGGGCAGAATGTCAAGAGCTCCTCTGCA
GCTTCAGGTGTGACTTGTCCATCTACTGTGGTCAAGAAGATGAAGGTGTCTCAGCAGAGCACACAA
AGCTCATACTGCCAAGTTGAAGGCTGCAAAGTCGATCTGTCCTCCGCAAGAGAATACCATCGCAAG
CACAAAGTTTGTGAAGCTCATTCTAAGGCACCAAAGGTTATTGTTTCTGGTCTGGAGCGCCGTTTT
TGCCAACAGTGTAGTCGGTTTCATGGTTTAGCTGAATTTGACCAGAAAAAGAAAAGTTGCCGCAGG
CGTCTATCTGATCATAATGCACGAAGAAGGAAACCGCAACAAGAGGCAATTTCATTTGGTTCATCA
AGGCTCGCCACGATGTTTTACGATGCAAGGCAGCAGACAGATATTTACTTTGGCCAATCTCCTTTT
GGCCAAGTGAGAAGCAATGCAATTTCTTCATGTGACAACCTGGGAGGCTTCAAATTTACAGAAGCA
AAACTCCCCTGGATGAAGCCAATGAAAACTATAGGCCTTGAGGATCTGAATTTCTCTACCCTGCAG
ATGCCAGGCAATGTTGTGTCGCATACGGTGCATCATCATGATTTTGATGGGCTCATACCATTCAAG
GGAAACACCCCGAAGGTCCTCAACCAAGGTGTGGATCCAGCCTGTGCCGTCGTGTCCTCCAACTCG
AATGGAGCCCCGGATCTTCGGCGTGCTCTCTCTCTTCTGTCAAGCGATTCCTGGGGCCCGGCCGAC
GTTCAGGCCGGCTCCCAGGTGCATCCCGGTGGAGTGATGCCGCCCCTCGCCGTTGCCGCCGCCACC
GTCACCGCCCCGACGAACCCTGTCAGTGTGATGCATGCTCTGCACCCGTCCACCGGAGGAGGAGGA
TTCTGGCAAGACGGCGACGACCCGCCTCCGCTCGATCATGCCTCGCAGGCTCAGGCGTTCATGCAT
CCTGGCAATGGCAGCAGCTCCGGCTATGGCCATCTGCACTGA

SEQ ID NO: 439, proteína - *Oryza sativa*

MASFGMNWNQKSPVFWDWENPAPFGPNTMENPKSIPHPEPRGVVVAAANHGSTNSSGGTFTSSSEL
ANGSSKSSLSASFDSSSKLGNSLEFRFASVKGHGKNMCKDGEAGRVEDSGTSPAVAVSHGEPVIGL
KLGKRTYFENVCGGQNVKSSSAASGVTCPSTVVKKMKVSQQSTQSSYCQVEGCKVDLSSAREYHRK
HKVCEAHSKAPKVIVSGLERRFCQQCSRFHGLAEFDQKKKSCRRRLSDHNARRRKPQQEAISFGSS
RLATMFYDARQQTDIYFGQSPFGQVRSNAISSCDNLGGFKFTEAKLPWMKPMKTIGLEDLNFSTLQ
MPGNVVSHTVHHHDFDGLIPFKGNTPKVLNQGVDPACAVVSSNSNGAPDLRRALSLLSSDSWGPAD
VQAGSQVHPGGVMPPLAVAAATVTAPTNPVSVMHALHPSTGGGGFWQDGDDPPPLDHASQAQAFMH
PGNGSSSGYGHLH

SEQ ID NO: 440, DNA - *Populus spp.*

ATGAGTTCTGTCTCACTGATGGAGTGGAATGGCAAACCCCACTTGCAGTGGGACTGGGAAAACCTG
ATAATGTTCAATGCAATAACAAATGAAAATTCAAAGCAGTTAAGCCTAACAGATTTGGAAACTGAT
GGAGAAAAAGGAAATGATTCTGGGTTTTTCTATTCATCTGGGAGTGTAAGCAGAAGCCGTGGATCT
ATCTCTGACTTAGAACTTGCTTCTTTCTCAAAGAGCTCGAAGTCAGCTTCCACCAATTCTTCATCA
GCTGGGGAAGTTAAAACATCCAAATTCACTTTGGGGGCCTCTAAAACAAATCCATCAGATTACAAT
AAGGAAGAACTTGTGAAGGCTAAGGCAACTGATACTTCTCCCACGCTTGAAGCTTCAGTTGGTTCA
GGTGACCAGCTGCTTGGTTTGAAGCTTGGTAAAAGAACTTACTTTGAAGATGCTTGTGCTGGGAGC
AATGCTCAGTCTTCATCAATATCTACGATTCCTGTGCCTTCTTTTACTCCAGCAAAGAAATTGAAG

FIGURA 32 CONT.

TCCAGTAATCATAGTCAGCGTGCTCCACGCTGTCAAGTAGAAGGCTGTAACCTTGACCTCTCATCA
GCTAAAGATTACCATCGCAAACATAGAGTTTGTGAAAGCCATTCAAAGTGCCCAAAGGTCATTGTA
GCTGGTTTGGAACGCAGGTTTTGCCAGCAGTGTAGCAGGTTCCATGGTCTGTCAGAGTTTGATGAA
AAGAAGAAAAGCTGCCGCAGGCGACTTTCTGATCACAATGCAAGACGCCGCAAACAGCCAGGATCT
GTCCATTTGAATCCTTCAAGACTTTCTTCATCATTATATGATGAGAGGCAACAGATGAGTCATGTT
TGGGACAAGGCGCCACTTGTTCATTCCAGGCCCAATGCAAATTTGACATGGGAAAGCACATCCACC
TCCAAGTTCACAATAACAAAAGAGTATATAGCAAAGCCTGCAGAAATAGGTGGTAGTGATAGGAGG
CTCCACTTGCCTGGCATTGATCTGACAAATAGCATTGCTATTCAGCACCATCATAAGTCTAATGGC
TTCTTACCATCCAAGGCCAAGGGCACTGCAGGTGAGGTTCTCAACCAAGGTTTGTTTTGTCTTCAT
TTTGAGAAACAAGCCTCACAATTATTGCATTGA

SEQ ID NO: 441, proteína - *Populus spp.*

MSSVSLMEWNGKPHLQWDWENLIMFNAITNENSKQLSLTDLETDGEKGNDSGFFYSSGSVSRSRGS
ISDLELASFSKSSKSASTNSSSAGEVKTSKFTLGASKTNPSDYNKEELVKAKATDTSPTLEASVGS
GDQLLGLKLGKRTYFEDACAGSNAQSSSISTIPVPSFTPAKKLKSSNHSQRAPRCQVEGCNLDLSS
AKDYHRKHRVCESHSKCPKVIVAGLERRFCQQCSRFHGLSEFDEKKKSCRRRLSDHNARRRKQPGS
VHLNPSRLSSSLYDERQQMSHVWDKAPLVHSRPNANLTWESTSTSKFTITKEYIAKPAEIGGSDRR
LHLPGIDLTNSIAIQHHHKSNGFLPSKAKGTAGEVLNQGLFCLHFEKQASQLLH

SEQ ID NO: 442, DNA - *Populus spp.*

ATGGAGTGGAATGGCAAACCCCACTTACAGTGGGACTGGGAGAGCCTGATAATGTTCAATGGAATA
ACAACTGAAAATTCTAAGCAGTTAAGCCCAACAGATTTGGAAACTGATGGAGAAAAAGGAACCGAC
TCTGGGTTTTTCTATTCATCTGGGAGTGCAAGCAGAAGCGGTGGTTCTAGCTCTGATTTGGAACTT
GCTTCTTTCTCAAAGTGCTCGAAGTCAGCTTCCATCAATTCTTCATCAGCTGGGGAAGTTAAAACA
TCCAAATTCACTTTGGAGGCCTCTAAAGCAAATCCATCTGATTACAATAAGAAAGAAATTGGAAAG
GCTAAGACAGCTAGTATGTCTTCCACAATTGAGGCTGCAGGTGGTTCAGGTGACCAGCTGCTTGGT
TTGAAGCTTGGTAAACGAATATACTTTGAAGATGCTTGTGCTGGCAACAATGTTCAGTCGTCATCA
TTTTCTACAGTTCCTGTGCCCTCTTTTACTTCAGCAAAGAAATTGAAGTCCACTATTCAGAGTCAG
CGTGCTCCATGCTGTCAAGTGGAAGGCTGTAACCTTGACCTCTCATCAGCTAAAGATTATCATCGC
AAACATAGAGTTTGTGAAAGCCATTCAAAGTGCCAGAAGGTCATTGTAGCTGGTTTGGAACGCAGG
TTTTGCCAGCAGTGTAGCAGTAAGATTATTCCTAGCTTTTATGATTGCATTGACAATTTGGGTTTT
GGATGTCAAAGGTTCCATGGCCTGTCAGAGTTCGATGAAAAGAAGAAAAGCTGTCGCAGGCGACTT
TCTGATCACAATGCAAGACGCCGCAAACAACCAGGATCGGTCCATTTAAATTCAAGAGTGTCTTCA
TCATTATATGATGAAAGGCAACAGATGAGTCTTGCTTGGGACAGGGCACCACTTGTTCATGCCAGG
CCTAATGCAAATTTGACGTGGGAAGGCACATACATCTCCAAGTTCACAATAACAAAAGATTATATA
GCAAAGCCTGCAGAAATAGGTGGTAATGATGGGCAGTTTCACTTGCCTGGCTTTGATCTGACAAAT
GGCATTGATATTCAGCACCATCATAAGTCTAATAGCTCCTTACCATCTAAAGGTAAGGGCACTGCA
GCTGAGATTCTCAACCAAGGTTTGGAAGAATACATCATTCCTTCCAAAGCGGAAGCAGCACCAGAA
TCTCATCGTGCTCTCTCTCTTCTGTCAAACAATTCATGGGGTTCACGTGAGCCGCAATCTATTTCA
TTTGAACAACCCGTGCATACAAATCACACCACTCAGTCTGTGCTGCAAGTTATACCCCAAAACTCA
CCACTTGCTTCATCAGAGTATTGGAGGACTGAGCAACCGTCAACTGACTCTCAAGTGCATACCTTG
ACGTCTCACTGA

SEQ ID NO: 443, proteína - *Populus spp.*

MEWNGKPHLQWDWESLIMFNGITTENSKQLSPTDLETDGEKGTDSGFFYSSGSASRSGGSSSDLEL
ASFSKCSKSASINSSSAGEVKTSKFTLEASKANPSDYNKKEIGKAKTASMSSTIEAAGGSGDQLLG
LKLGKRIYFEDACAGNNVQSSSFSTVPVPSFTSAKKLKSTIQSQRAPCCQVEGCNLDLSSAKDYHR
KHRVCESHSKCQKVIVAGLERRFCQQCSSKIIPSFYDCIDNLGFGCQRFHGLSEFDEKKKSCRRRL

FIGURA 32 CONT.

SDHNARRRKQPGSVHLNSRVSSSLYDERQQMSLAWDRAPLVHARPNANLTWEGTYISKFTITKDYI
AKPAEIGGNDGQFHLPGFDLTNGIDIQHHHKSNSSLPSKGKGTAAEILNQGLEEYIIPSKAEAAPE
SHRALSLLSNNSWGSREPQSISFEQPVHTNHTTQSVLQVIPQNSPLASSEYWRTEQPSTDSQVHTL
TSH

SEQ ID NO: 444, DNA - *Populus spp.*

ATGGACTGGAACTCCACTGCATCTGAGGGGAACTGGGAGAACATAGCAGGGTTAAGTGCCAAGGCT
AGTGACATTCCAAAACAAGTACCATTGGCATACCATGAGACAGAGGGAGATGGAGCAATTGATAAA
GCAATAACTTATTCGTCTGGTGGTGGTTTATCAAGCTCTGATTTGTGGCATTGCTCTTCATCCAAG
AGCTCAGTTTCACCCTCTGTCGACTCTTCATTGAAGGGGGGGATCAAGACTTATAGCACTGCAGAT
GGTTTTCCTGGAGTTATCATTAGAAAGGACTTGACTAGGGTAGAAAGTACTGAGAATATTCCATCT
CTGGGTGCTTCTGCTAGCTCTGGTGAACCAGTAATTGGCTTGAAGCTTGGTAAACGGATGTACTTT
GAAGACATTTGTACTACCAGCACTGCCAAATCATCGTCTTCATCTATTGTTTCCACTTCCTGTACT
GCTACAACAAAGAGATCTAGGGCATCATATCCAAGTACACAGTCTCCACGTTGCCAAGTGGAAGGA
TGCAACCTTGATCTCAAGTCAGCTAAAGATTACCATCGCAGGCATAGAATCTGTGAAAAACATTCC
AAAAGCCCAAAGGTTATTGTTGCTGGCATGGAACGCCGATTTTGCCAACAGTGCAGCAGGTTCCAT
GAATTATCAGAGTTTGATGACAAGAAGCGAAGCTGTCGTAGACGTCTCTCTGATCACAATGCAAGG
CGACGGAGGCCACAACCTGAGGCAATCCGGTTCAATTCAGCAAGGCCATCATCATCACCATCATCA
TTTTATGGTTGTCATTTCAATTTCAAAGATGGTAGACAGCAGATGAATGTTGTATTGAATAGGGTG
CCTTTCCTGGCAGGAAATTCAACTTGGCAAAGCGAGTGCGGCTTTAAAGTCACTCACGCCGGAGAT
TTTTTTATTAGGCCTGCAAAAGCAGGGGGCATCCATAGGCAGCTTAGCTATCCCTGCAATGAAGTG
CCAGATGCAGCTTCAACAATACCAACAGACTCGGATAAAATTATCTCTTTTGAGAGAAGTACACCC
CAGGTCTTTAGTCAAGGTTTCGAGGGGTCTGCATTAACTTCTAACATTGAGGTAGCACCAGATCTT
CGGCGTGCTCTCTCTCTTCTGTCAACCACTTCTTGGGGCTCGAATGAGCATGGATCCACCTCTCTG
GATCAACTGATGCTTGCAAACCAAACAAGCATGACACAGCCAATGATAAATGCTGAACTCCAAAAT
TGTCCAGTTGCTTCATCAGAAAATGCACGGATGGAACAAGCATCCCTTGAGTCTCGGGTGCATTCT
TTGGATTTACATGACAATGGAAACGTCCAGTTGCAAGAGTTTCAGCTGCTCAAAGCACCCTATGCT
ACCGGCTGCTTTTATTCCAATCAATTCAGCTAA

SEQ ID NO: 445, proteína - *Populus spp.*

MDWNSTASEGNWENIAGLSAKASDIPKQVPLAYHETEGDGAIDKAITYSSGGGLSSSDLWHCSSSK
SSVSPSVDSSLKGGIKTYSTADGFPGVIIRKDLTRVESTENIPSLGASASSGEPVIGLKLGKRMYF
EDICTTSTAKSSSSSIVSTSCTATTKRSRASYPSTQSPRCQVEGCNLDLKSAKDYHRRHRICEKHS
KSPKVIVAGMERRFCQQCSRFHELSEFDDKKRSCRRRLSDHNARRRRPQPEAIRFNSARPSSSPSS
FYGCHFNFKDGRQQMNVVLNRVPFLAGNSTWQSECGFKVTHAGDFFIRPAKAGGIHRQLSYPCNEV
PDAASTIPTDSDKIISFERSTPQVFSQGFEGSALTSNIEVAPDLRRALSLLSTTSWGSNEHGSTSL
DQLMLANQTSMTQPMINAELQNCPVASSENARMEQASLESRVHSLDLHDNGNVQLQEFQLLKAPYA
TGCFYSNQFS

SEQ ID NO: 446, DNA - *Brassica rapa*

ATGGAGTGTAATGCAAAGCCATCATTGCAGTGGGAGTGGGATAATCTAATATCTTTTGGTACTTCA
TCAGCTGAAATTCCTAAAAAGCAACGACCAATGGATTGGGAAAATGATGGGTTTGATTGCACCACT
TTTTACTCGTCCAGCTTTGCTACAGAAGCAGCTTATGGTGGTGGTAGTTCAGGTTCTGATCTAGCT
CATGCTTTCTCTAAAAGCTCAAAGTCAACTTCCATAAGCTCTTCATCAGCTGAAGTGAGAACATAC
AATTTCACATCAGAAGCTGGTGAAAGTGTTCCTCCTGGAGAATTGGGGAGCAGTGAAGAGTTTGCA
ATGGGAATTGATGCTTCTCCAAGTCTTGAACTCTCCTTCGGCTCTGGTGATCCGGTTCTTGGTTTG
AAGCTTGGTAAGAGGACATACTTTGAAGACTTTTGGGAAGTGGAGAATGCTAAAGGTTCAGCACTT
CCGGTGAGCCTGGCGTCATCTTCTGCTTCTCCGGTGAAGAAATCCAAAACACTCTCTCAGAAATTA

FIGURA 32 CONT.

CAAACTCCTCACTGCCAAGTTGAAGGCTGTAACCTCGATCTCTCCTCAGCTAAGGACTATCATAGG
AAACACAGGATTTGTGAAAACCATTCAAAGTTCCCTAAAGTCGTTGTCAGTGGCGTAGAGCGCCGG
TTCTGCCAACAATGTAGCAGGTAGAAAAGCTAGTCTTAGAGATATTGGTTTATACAAAAGCCTTAA
TGACTTTTTTTAATTGTTTTTGGTCTTCAAAGGTTTCACTGTCTCTCTGAGTTTGATGAGAAGAAA
CGTAGCTGTCGCCGGCGTCTCTCAGATCACAATGCAAGACGTCGCAAGCCAAATCCCGGGAGGACA
TATGGTAATATACTAGATTCACTTCTCTCTAATCCTATATATTAAATCTAAATTTTTAGTTATATA
ATTAGTAGGTTTCCCTTTTCTAAGATAAATATCATTCTGTATGTGCAGATGGGAAGCAACAGATGG
ATTTTGTATGGAACAGATTTGCACTTATCCATCCAAGAAGTGAAGAAAATTTTCTGTGGCCAAATC
CGAAGCCTGTATCATCAAGAGGGTTATTGCAGGAACCTGCAAAGACCGAGATGCCCAATAAGCTTT
TCACCGAGCATTGTGGATTTGGATTGTTGGACCCCAAAACGAAAACCACCAGAGCTGAGTTATTCA
GTAAAGGTTTGTGTTTTTCTTTGCTGTTTAATATATTTTCTCTTGTTGTATTCATGATCTTGATGA
TTGCTTGTTTTATGCAGATTTGTTAGAAAAGGTCACAATCTCTTCTTCACACATGGGTGCTTCTCA
AGATCTTGATGGTGCTCTCTCTCTTCTGTCAAATTCAACACCATGGGTTTCCTCGAACCAGCCAAC
ACGGTTTTCCCTTAACCACCATTCCACAAGCAACCTCCAACCCGTGGTTCACGGGTCTGTGACTCA
ACTCAGTTCAGTGTCCAGCTACTGGCAGCCGGACCCACCAGCAGCTGAGGGCTCAACCGCTTTGAC
TAGGAATGGGGTAGGCCAGTTTAATGAGAACTACAACTTGAATCAGTTTTATAACTGAAAGTGTGT
TGCTTTTAAAATCATAATAAGATATTTTGTGTAAGGATCAGGTGAGCCTAGTGGTAACGTTTAAAT
AGGAGCTGTGAAACTTGCTAAAGACATCAACTCCTCTCGTCTTTCTTTTGTCCATTGATTTTCTTG
GGTTAGGGAGAGTTGTGACAGTTAAGTATTATAAATCATCTTGTCAAATTATTTATACATTAGAAT
TTTCAGATATTGATGGTGTAGCAACTGAAAAATTGATCAAACGCATCATAATATACGTATGATAAG
GTTATTTGCATCTGAACTTTTTTTATATTCGCATTAGTCAGCATTTTGTGTTATACTAGATTTGGA
TCCGTGCGTTGCAACAGGTTTTATTTGATATTACCATCCATTGATTTTTTTTTTTTTTTGCAAACA
AATAATTTGGTAATTCCATTAACCGCCACAGTCGTTTTCTGAGCAAAACCATGTTAAAATCT

SEQ ID NO: 447, proteína - *Brassica rapa*

MECNAKPSLQWEWDNLISFGTSSAEIPKKQRPMDWENDGFDCTTFYSSSFATEAAYGGGSSGSDLA
HAFSKSSKSTSISSSSAEVRTYNFTSEAGESVPPGELGSSEEFAMGIDASPSLELSFGSGDPVLGL
KLGKRTYFEDFWEVENAKGSALPVSLASSSASPVKKSKTLSQKLQTPHCQVEGCNLDLSSAKDYHR
KHRICENHSKFPKVVVSGVERRFCQQCSRKSSRYWFIQKPLFLIVFGLQRFHCLSEFDEKKRSCRR
RLSDHNARRRKPNPGRTY

SEQ ID NO: 448, DNA - *Zea mays*

AGCCGCTTCGGCAGGCAAGTAGGACGAGGAGCAGCAGCTCAGATTTCGGATCCTGGCACTGGCAAC
TGGCATGGGCTCGTTTGGGATGGACATGGACTGGAACCAGAAGGCCTCCGTGCTGCTGTGGGACTG
GGAGAACCTGCCGCCCGCAGCCGCAAACGGGAGCGAGAGCCACAGGACGACCGCTGCCGCGCCTCA
GTTCGCGGGCCTTGAGGCCACAGGGCATGAACCGGCGCCTTCTTCCGTATCCTCGTCAATCCATTC
TCTGCCCAGTGCCAAGGGGAACAAGAACAAGAACAACATGGAGCTCATCAGTTTCGCACCTGCCAA
AGCGCCCGACAAGGACACTGGTTCGGTGCTCAGCAGCGGAGAGCCGGTGGTGCTAGGCCTGAAGCT
TGGCAAGAGAACGTATTTCGAAGATGGTTGTGGACTAGGGCAGAGCGGCAAGAGTTCGGCGGCGTT
AGGCACTGCCAGTGCAGCGACCCCTCCGGGTCCTGCGAAGAAGGCGAAGGCGGCGGCGGCGGCTCC
AACCGCGCAGCAGCAGAAATCGTACTGCCAGGTTGAAGGCTGCAGGACCGATCTGTCCTCTGCTAA
AGACTATCATCGCAAGCACAGAGTCTGCGAGCCCCATTCCAAGGCGCCCAAGGTGGTCGTCGCTGG
CCTGGAGCGGCGCTTCTGCCAGCAGTGCAGCCGGTTCCATGGACTGGCCGAGTTCGACCAGAAGAA
GAAGAGCTGCCGCAGGCGACTCAACGACCACAACGCGCGCAGGCGGAAGCCCCAGCCCGAAGCGCT
CCCTTTCGGCTCATCGAGGCTGTCGGCGATGTTCTATGACACGAGGCGCCAGGCGAGCCTCCTGTT
TGGTGAAGGTCCCTACGGTCAGATGAGAAGCTGTGCGAGTTCTTGGTGGGATAGCCCAGGAGCAGG
AGGAGGAGGCTTCAAATTCGCGGAGACGAGAGGAGCTCCTTGGTTGAAGCCGGCGAGAGCTGCTGC
TGATGGGCTGCTGCCTTTGTCAGGTCAGCAGCAGCAGCAGGTGTGGAATGACTCGATTCCGCATGT

FIGURA 32 CONT.

TGCCCATCAGGAGGATTTCGATGGGTTCACGGTCACGGCGTTCAAAGGAATCAGTCCAAAGACCCT
TGATCATCAAGGCGCAGGCGTTGAAGCCTGTGCGGCCGTCTCCAGCCCGGACGGCGCCCCGGACCT
TCAGCGTGCTCTCTCTCTTCTGTCAAACAGTCCGGCCGGTGGCGGCGGCACCAACACCAACCACCC
GCCGCCGCCGCCGGCGGCTGAGCTGCACCACCACCGCGCTGCCCCGAGCAGCTGCCTCGCCGCCGG
CTCCGTCTCCGTGCTGCAGGAGGCCTCGTCGCCGGGGCTGTGGCAGGACGGCGGCGGCAGCGCGGC
CCTGGGCCACCACGCGCACGCGCGGCTCCAGGCTCTCGACGCCGCCATGGCGACGGCGCAGGAGCT
CCTCCAGCTCCCCAGGCCGCCGCCATCGTACGGCGGCTCCTCGTCCCACTACGACCTGATGCGCTG
ATGCTTCCTTCCTTGTGGATTGTGGTGGAGAACTGGAGATGGAGCGGGAGCGCCGCCGTGCCCCCG
GTTGGCGCGTGGAAATCAGTCAGACGTGGCTTCTGTGGCGTCGTCTCTGCAACGCGGGAGCGGGAT
TTTGCTTGTGTTCTGAACGAACAGAAAAAAAAAA

SEQ ID NO: 449, proteína - *Zea mays*

MGSFGMDMDWNQKASVLLWDWENLPPAAANGSESHRTTAAAPQFAGLEATGHEPAPSSVSSSIHSL
PSAKGNKNKNNMELISFAPAKAPDKDTGSVLSSGEPVVLGLKLGKRTYFEDGCGLGQSGKSSAALG
TASAATPPGPAKKAKAAAAAPTAQQQKSYCQVEGCRTDLSSAKDYHRKHRVCEPHSKAPKVVVAGL
ERRFCQQCSRFHGLAEFDQKKKSCRRRLNDHNARRRKPQPEALPFGSSRLSAMFYDTRRQASLLFG
EGPYGQMRSCASSWWDSPGAGGGGFKFAETRGAPWLKPARAAADGLLPLSGQQQQQVWNDSIPHVA
HQEDFDGFTVTAFKGISPKTLDHQGAGVEACAAVSSPDGAPDLQRALSLLSNSPAGGGTNTNHPP
PPPAAELHHHRAAPSSCLAAGSVSVLQEASSPGLWQDGGGSAALGHHAHARLQALDAAMATAQELL
QLPRPPPSYGGSSSHYDLMR

SEQ ID NO: 450, DNA - *Zea mays*

ATGGGTTCTTTTGGGATGAACTGGAATCAGAAGGACCCCATGGTGTGGGATTGGGAACATCTAGTA
CCGTCTGTCTCAAATGCAGTTACAAGGCACGGATCTGCTAATTCATCTGGTGGTACTCTTACTTCT
AACTCAGAGCTAGGGCATGGTTCATCCAAGAGCTCTATTTCAGCGTCCATTGATTCACCCTCTGGA
GTAGGGAACAGCTTAGAGTTCAACTTCGCCGCTGTTGAGAGGCATGTTAAGAACACGGGCACGAAC
GGCAGAGTCGATGACTCGGGGAATTCTCCATCGTCAATGATAGCTTTCAACCAAGGAGAGCCATTA
ATCAGCCTGAAACTTGGGAAGAGGGCTTACTTCGAAAACGCCTGCGGAGGACAGGATGCCAAGGTT
TCTGCAGCTTCAGACGTTACTTCTGCAGCCAGCGTGGTCAAGAAGACTAAGGTGTCTCAGCAGAAT
GCAAAGAACTGGTACTGTCAGGTTGAAGGGTGCAAAGTTGACCTGTCTTCTGCTAAAGATTACAAT
CGCAAGCACAAGGTCTGTGTAGTCCATTCTAAAGCTACCAAGGTGGTTGTTGCTGGTCTAGAGCGT
CGGTTTTGTCAACAGTGTAGCCGTTTTCATGGTTTAGCGGAGTTTGATCAGAACAAACGAAGCTGT
CGTAGGCGTCTGATGCATCATAATGCACGGAGGAGGAAACCTCAGGCAGACACAATTTCATTCAAT
TCATCGACAATGTTTTATGATACAAGGCAGCGGACAAATCTTTTCTTTAGTCAACCACTTTATGGC
CAAGTGAGGAGCAATGCAGGGTCTTCATGGGATAACTTGGGAGGCTTAAAATTCATGGAGACGAAA
CATCCGCCAGTGCATCCAACAAAAACAGCATCCCCTGATGAACTGCATTTCTCAGCCCTCCAGATA
ACTAGTGCTGCGGCTCACACCGGACATCATCATGATCTCGATGGGTTCATGGCGTTCAAGGGAACC
AGCACAAAGGTCCTTAACCAAGGCGTGGAGGCTTGGGCGGCCGCTTCCAGCTCGAACAACGGAGGC
CCAGAAGGTGGGCGTGCTCTCTCTCTTCTGTCAGACGGCTCGTGGGGCTCGAGTTCAGCCGTCATC
CAGCAGCCCACATCTCACGCGGACGCCGGTGCATTGCTGCCGCCCCTCGCCACCGTTGCCGTCTCC
AACGCCGCCGCCGCCGCCGGGCATCCTCTGGACCCGTCCCCGGGAAGGTTCTGGCCGCAAGACGAT
CATCCCCCGCTCGTCGACGGACCCGCCACGCAGATTCCGGAGCTGGCGCACCTCCGGATATGGTGA

SEQ ID NO: 451, proteína - *Zea mays*

MGSFGMNWNQKDPMVWDWEHLVPSVSNAVTRHGSANSSGGTLTSNSELGHGSSKSSISASIDSPSG
VGNSLEFNFAAVERHVKNTGTNGRVDDSGNSPSSMIAFNQGEPLISLKLGKRAYFENACGGQDAKV
SAASDVTSAASVVKKTKVSQQNAKNWYCQVEGCKVDLSSAKDYNRKHKVCVVHSKATKVVVAGLER
RFCQQCSRFHGLAEFDQNKRSCRRRLMHHNARRRKPQADTISFNSSTMFYDTRQRTNLFFSQPLYG

FIGURA 32 CONT.

QVRSNAGSSWDNLGGLKFMETKHPPVHPTKTASPDELHFSALQITSAAAHTGHHHDLDGFMAFKGT
STKVLNQGVEAWAAASSSNNGGPEGGRALSLLSDGSWGSSSAVIQQPTSHADAGALLPPLATVAVS
NAAAAAGHPLDPSPGRFWPQDDHPPLVDGPATQIPELAHLRIW

SEQ ID NO: 452, DNA - *Triticum aestivum*

ATGGGCTCGTTCGGGATGGAGTGGAACCAGAGGAGCTCGGTGCTGTGGGACTGGGAGAATTTCCCG
CCGATAGGCGAGAACCCCAAGAACGCGATGCAGGCCGATCCAAGATTTGCCGCCGTTGCGGCTACC
ATGGGGAATGAACCGCTCCATTCTTCTGGCGGTAGCGGCACCTTCTCTTCCAGCTCAGAGATGGGG
TATGGCTCTTCCAAGAGCTCCATGTCCGCGTCGATCGATTCTTCGAACAGGGCTGGGAACAACATG
GAGTTCAGATTTGCGCCTGTCAAAAACCCTGATAGGAACACGAGCAAGAACACCGAGCTGGGTAAA
GTTGATAACACGAGAACTGGAACATCTCCGTCGCCTGTGGTGGCAGTGAGCAGTGGAGAGCCGGTG
ATCGGCCTGAAGCTTGGCAAGAGAACTTACTTCGAGGATGTCTGTGGAGGGCAGAATGTCAAGAGC
TCGCCATCGGGTGCTGCGAGCGCGCCAAACAAATCTCCTGCTTTGGGCAAGAAGGCAAAGGCGGAA
CAACAGAAGCCACATAACTCGTACTGTCAGGTTGAAGGCTGCAAAGTCGACCTCTCTTCTGTTAAA
GATTACCATCGAAAGCACAGAGTCTGTGAACTTCACTCTAAGGCTCCGAAAGTTGTTGTCGCTGGT
CTGGAGCGACGCTTTTGCCAACAGTGCAGCCGGTTTCATGCTTTAGCTGAGTTTGACCAGAAAAAG
CGAAGCTGCCGTAGGCGTCTCAATGATCATAATTCCCGCAGGCGGAAGCCACAGCCAGAAGCAATT
TCTTTCAGTTCATCAAGGATGTCTACGATGTTTTATGATGCAAGGCAACAGCCAAATTTCCTATTT
GGTCAGGCTCCTTATGTTCAAATGAGAAGCTGTGGAAGTTCTTCATGGGATGACCCAGGAGGCTTC
AAAGTTACACACACAAAAGCTCCTTGGTTAAAACCAACAACTGCTGCAGGTGTTCATGGGATACAT
TTATCTAGTCAGCAGATGTCGGACAATATTATGCCACATGGTGCACATCATGGTTTCGATGGGTTC
ATGGCATTCAAGGGAACTTGTACAAAGTTCCCTAATCAAGGTGTCCAAGCTTCTGCTGTTGCTTCC
GACTCCAGTGGAGCCCCGGATCTTCAGCATGCTCTCTCTCTTCTGTCAAGCAACCCAGTGGGTGCT
GCCAACCTCCAGCCAAGTCCCCAGATGCACTCTGGGGTGGCAGCCATTGCCGGCGCCCCCAACCCC
GCGATGCACGCGCTGGGATCATCGACGGGGCTCTGGCTAGACGGCAGCCAGCCCCTCGATGATCAC
CCGCGGTTCCAGGTCTTCGAGCGCTTGGGGGACCATGACAGCGAGCTCCAGCTCCCAAAGCCTTCC
TACGACCATGCCTCGCACTTCGACCGGATGCACTGA

SEQ ID NO: 453, proteína - *Triticum aestivum*

MGSFGMEWNQRSSVLWDWENFPPIGENPKNAMQADPRFAAVAATMGNEPLHSSGGSGTFSSSSEMG
YGSSKSSMSASIDSSNRAGNNMEFRFAPVKNPDRNTSKNTELGKVDNTRTGTSPSPVVAVSSGEPV
IGLKLGKRTYFEDVCGGQNVKSSPSGAASAPNKSPALGKKAKAEQQKPHNSYCQVEGCKVDLSSVK
DYHRKHRVCELHSKAPKVVVAGLERRFCQQCSRFHALAEFDQKKRSCRRRLNDHNSRRRKPQPEAI
SFSSSRMSTMFYDARQQPNFLFGQAPYVQMRSCGSSSWDDPGGFKVTHTKAPWLKPTTAAGVHGIH
LSSQQMSDNIMPHGAHHGFDGFMAFKGTCTKFPNQGVQASAVASDSSGAPDLQHALSLLSSNPVGA
ANLQPSPQMHSGVAAIAGAPNPAMHALGSSTGLWLDGSQPLDDHPRFQVFERLGDHDSELQLPKPS
YDHASHFDRMH

SEQ ID NO: 454, DNA - *Vitis vinifera*

ATGGATGCTGTTGTTTTGACAGGTTTGGAAGAATCTGGGGTAAGAGTAAGGGTGGTTAATTTTATA
CAGGGCGGGCGAGAAATGGCATTTTTGTTGGAATCTGTGCAAGAAGAGGACTTGGGTAGAAACATC
ATCGGTTCTACTTGGTCGACTTGGGGAAGAAAGGACAATCGATATCCTTCTCTAAAAGGGGATAAG
GAATTTCCTTTCTGTGGCATGTGGGACTGGGAGAACCTTGACATTTTCAATGCAAAATCAACTGAA
ATTCCAAAAAAGTTACAATCAGACTGGGAAATTGAAGGAGAAGGAGGAATTGACACTGGTTCTTTC
TATTCATCTGGATGTGGTGCTGGTAGTGGTGGTTCTGGCTCTGATTTGGGACATGCATATTCATCA
AAGAGCTCAAAATCAGCTTCAATTGATTCTTCATCAAAGGTGGGAATAAAGACATCCAACTTCACT
TTTGAGGCTTTCCAAGATTTTTCTCAAGATTTTAACAAGAAGAGAGATTTGGAGAGGGCTGAGCCA
ACTGGAAGTTCTCCAAACCTTGAAGCCTCGATTGGCTCTGGTGAACCTTTAATTGGTCTAAAGCTT

FIGURA 32 CONT.

GGTAAAAGGACATACTTCGAGGATGTTTCTGCAGGGGGCAATGCTAAGGGCTCAGTGTTTTCTGTG
ATTCCCATATCATCTGATACCACAGCAAAGAGATCCAGGTTATCTTGTCAGAATGCACATGTCTCG
CGCTGCCAAGTTGAAGGATGCAACCTTGACCTTTCAACAGCCAAAGATTACCATCGCAAACATAGA
GTTTGTGAAAGTCATACCAAATGCCCAAAGGTCATTGTAGGTGGTCTGGAGCGCCGTTTTTGCCAA
CAGTGTAGCAGGTTCCATAGTTTGTCAGAGTTTGATGAAAAGAAGAGAAGCTGTCGCAGGCGTCTT
TCTGATCACAATGCACGCCGGCGCAAACCACAACCAGAAGGGATCCAGTCAAATTCGGCGAGGCTT
TCATCATTCTATGGTGGGAAGCAGCAAATGAGTCTTGTGTTGAGCTCGGTCCCCATTGTTCATACA
AGGCCTGCTGTAAATCCTTCATGGGAAGGCACATGCAGCAAGTTCACACAAACAAAAGGGTTAGTA
AGGCCTGGTAACACAGGAGGTATCGACCGGCAGCTGCATTTGCCTGGCAGTGAGCTGCAGAATGGC
ATTAGTACGCTTTGTCATGATTCCAGTAGGCTTTTGCACTCTAAGGGCGCAATAGCTGAGGTTCTC
AATCAAGTGTCCCAAGGTATGTTACATTCAACCTCAGAACACTGGAGGACCGAGCTTCCCTCAACT
GATTCCCGGGTGCATATCTTGCCTTCACAAGGCGATGGTAACACCCACTTTCAAGAGTTCAAGCTG
TTCAAAGCGCCATACGAGTCTGGCTTTTCTTCTGGATATTCAGAGTATAGGGAATTGAGATATCCG
ATTCCTAGGAGTAACGCCCGTGTCATGAAGCAAACATAG

SEQ ID NO: 455, proteína - *Vitis vinifera*
MDAVVLTGLEESGVRVRVVNFIQGGREMAFLLESVQEEDLGRNIIGSTWSTWGRKDNRYPSLKGDK
EFPFCGMWDWENLDIFNAKSTEIPKKLQSDWEIEGEGGIDTGSFYSSGCGAGSGGSGSDLGHAYSS
KSSKSASIDSSSKVGIKTSNFTFEAFQDFSQDFNKKRDLERAEPTGSSPNLEASIGSGEPLIGLKL
GKRTYFEDVSAGGNAKGSVFSVIPISSDTTAKRSRLSCQNAHVSRCQVEGCNLDLSTAKDYHRKHR
VCESHTKCPKVIVGGLERRFCQQCSRFHSLSEFDEKKRSCRRRLSDHNARRRKPQPEGIQSNSARL
SSFYGGKQQMSLVLSSVPIVHTRPAVNPSWEGTCSKFTQTKGLVRPGNTGGIDRQLHLPGSELQNG
ISTLCHDSSRLLHSKGAIAEVLNQVSQGMLHSTSEHWRTELPSTDSRVHILPSQGDGNTHFQEFKL
FKAPYESGFSSGYSEYRELRYPIPRSNARVMKQT

SEQ ID NO: 456, proteína - domínio sbp da SEQ ID NO: 428
RCQIDGCELDLSSAKGYHRKHKVCEKHSKCPKVSVSGLERRFCQQCSRFHAVSEFDEKKRSCRKRL
SHHNARRRKPQGV

SEQ ID NO: 457, proteína - domínio sbp da SEQ ID NO: 433
RCQIDGCELDLSSSKDYHRKHRVCETHSKCPKVVVSGLERRFCQQCSRFHAVSEFDEKKRSCRKRL
SHHNARRRKPQGV

SEQ ID NO: 458, proteína - domínio sbp da SEQ ID NO: 435
HCQVEGCNLDLSSAKDYHRKHRICENHSKFPKVVVSGVERRFCQQCSRFHCLSEFDEKKRSCRRRL
SDHNARRRKPNPG

SEQ ID NO: 459, proteína - domínio sbp da SEQ ID NO: 437
HCQVEGCNVDLSSAKPYHRKHRVCEPHSKTLKVIVAGLERRFCQQCSRFHGLAEFDQKKRSCRRRL
HDHNARRRKPQPE

SEQ ID NO: 460, proteína - domínio sbp da SEQ ID NO: 439
YCQVEGCKVDLSSAREYHRKHKVCEAHSKAPKVIVSGLERRFCQQCSRFHGLAEFDQKKKSCRRRL
SDHNARRRKPQQE

SEQ ID NO: 461, proteína - domínio sbp da SEQ ID NO: 441
TPAKKLKSSNHSQRAPRCQVEGCNLDLSSAKDYHRKHRVCESHSKCPKVIVAGLERRFCQQCSRFH
GLSEFDEKKKSCRRRLSDHNARRRKQPGSVHLNPSRLSSSLY

FIGURA 32 CONT.

SEQ ID NO: 462, proteína - domínio sbp da SEQ ID NO: 443
CCQVEGCNLDLSSAKDYHRKHRVCESHSKCQKVIVAGLERRFCQQCSSKIIPSFYDCIDNLGFGCQ
RFHGLSEFDEKKKSCRRRLSDHNARRRKQPGS

SEQ ID NO: 463, proteína - domínio sbp da SEQ ID NO: 445
DLKSAKDYHRRHRICEKHSKSPKVIVAGMERRFCQQCSRFHELSEFDDKKRSCRRRLSDHNARRRR
PQPEAIRFNSARPSSSPSSFYGCHFNFKDGRQ

SEQ ID NO: 464, proteína - domínio sbp da SEQ ID NO: 447
TLSQKLQTPHCQVEGCNLDLSSAKDYHRKHRICENHSKFPKVVVSGVERRFCQQCSRKSSRYWFIQ
KPLFLIVFGLQRF

SEQ ID NO: 465, proteína - domínio sbp da SEQ ID NO: 449
SYCQVEGCRTDLSSAKDYHRKHRVCEPHSKAPKVVVAGLERRFCQQCSRFHGLAEFDQKKKSCRRR
LNDHNARRRKPQPEAL

SEQ ID NO: 466, proteína - domínio sbp da SEQ ID NO: 451
CQVEGCKVDLSSAKDYNRKHKVCVVHSKATKVVVAGLERRFCQQCSRFHGLAEFDQNKRSCRRRLM
HHNARRRKP

SEQ ID NO: 467, proteína - domínio sbp da SEQ ID NO: 453
YCQVEGCKVDLSSVKDYHRKHRVCELHSKAPKVVVAGLERRFCQQCSRFHALAEFDQKKRSCRRRL
NDHNSRRRKPQPEAISFSS

SEQ ID NO: 468, proteína - domínio sbp da SEQ ID NO: 455
RCQVEGCNLDLSTAKDYHRKHRVCESHTKCPKVIVGGLERRFCQQCSRFHSLSEFDEKKRSCRRRL
SDHNARRRKPQPEGI

SEQ ID NO: 469, proteína - motivo 1
WDWENL

SEQ ID NO: 470, proteína - motivo 2
SKSSQSTSINSS

SEQ ID NO: 471, proteína - motivo 3
ALSLLST

SEQ ID NO: 472, proteína - motivo 4
LKLGKR

SEQ ID NO: 473, DNA - iniciador 1
ggggacaagtttgtacaaaaaagcaggcttaaacaatggactgcaacatggtatct

SEQ ID NO: 474, DNA - iniciador 2
ggggaccactttgtacaagaaagctgggtcacattacttactcatctattttgg

SEQ ID NO: 475, DNA - sequência promotora
TNCGTACAA

FIGURA 32 CONT.

SEQ ID NO: 476, DNA - *Oryza sativa*

AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAATAT
AAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCACCTACT
TTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGTAATTAAGT
GGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGTTAAATTATTCG
AGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTTCTAGCTGAACTCAATGGGT
AAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTCTGAACGTATTGGCAAAGATTTAAA
CATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACATCATTAAGGACATGTCTTA
CTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATTTTTATTATTTATCTTTTTTCG
ATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAAT
ACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTAGGTAGCAATATC
TGAATTCAAGCACTCCACCATCACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTT
ACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAAAAAAGAATT
TTGCTCGTGCGCGAGCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCCACA
GAACAACCCACAAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAG
GCTTTGCGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGCATCCTCCTTCTCCCATCTATAA
ATTCCTCCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAAG
GACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTCGATCCATATCTTCCGGTCGAGTTCTTGGT
CGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCTCCCTTCGGTTGTTCTTGGATTT
ATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCC
TGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTCGGTTTGATTAGTAGT
ATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTAGGGATCGGAATCTTGCGATT
TTGTGAGTACCTTTTGTTTGAGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAGTAC
GGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTC
CCTATTGAACAAAAATAATCCAACTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTT
AAGCCTGTCCAAAATTTCGCAGCTGGCTTGTTTAGATACAGTAGTCCCCATCACGAAATTCATGGA
AACAGTTATAATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTT
TTTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATG
CTTTTATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGA
AGAACTTATCCGATTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATT
CATTTGGATTATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAA
CTGTCCTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGT
AGAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAATCG
GGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCTTGCCACT
TTCACCAGCAAAGTTC

SEQ ID NO: 477, DNA - *Oryza sativa*

GCTTGAGTCATAGGGAGAAAACAAATCGATCATATTTGACTCTTTTCCCTCCATCTCTCTTACCGG
CAAAAAAAGTAGTACTGGTTTATATGTAAAGTAAGATTCTTTAATTATGTGAGATCCGGCTTAATG
CTTTTCTTTTGTCACATATACTGCATTGCAACAATTGCCATATATTCACTTCTGCCATCCCATTAT
ATAGCAACTCAAGAATGGATTGATATATCCCCTATTACTAATCTAGACATGTTAAGGCTGAGTTGG
GCAGTCCATCTTCCCAACCCACCACCTTCGTTTTTCGCGCACATACTTTTCAAACTACTAAATGGT
GTGTTTTTTAAAAATATTTTCAATACAAAAGTTGCTTTAAAAAATTATATTGATCCATTTTTTTAA
AAAAAATAGCTAATACTTAATTAATCACGTGTTAAAAGACCGCTCCGTTTTGCGTGCAGGAGGGAT
AGGTTCACATCCTGCATTACCGAACACAGCCTAAATCTTGTTGTCTAGATTCGTAGTACTGGATAT
ATTAAATCATGTTCTAAGTTACTATATACTGAGATGAATAGAATAAGTAAAATTAGACCCACCTTA
AGTCTTGATGAAGTTACTACTAGCTGCGTTTGGGAGGACTTCCCAAAAAAAAAAGTATTAGCCATT
AGCACGTGATTAATTAAGTACTAGTTTAAAAAACTTAAAAAATAAATTAATATGATTCTCTTAAGT

FIGURA 32 CONT.

AACTCTCCTATAGAAAACTTTTACAAAATTACACCGTTTAATAGTTTGGAAAATATGTCAGTAAAA
AATAAGAGAGTAGAAGTTATGAAAGTTAGAAAAAGAATTGTTTTAGTAGTATACAGTTATAAACTA
TTCCCTCTGTTCTAAAACATAAGGGATTATGGATGGATTCGACATGTACCAGTACCATGAATCGAA
TCCAGACAAGTTTTTTATGCATATTTATTCTACTATAATATATCACATCTGCTCTAAATATCTTAT
ATTTCGAGGTGGAGACTGTCGCTATGTTTTTCTGCCCGTTGCTAAGCACACGCCACCCCCGATGCG
GGGACGCCTCTGGCCTTCTTGCCACGATAATTGAATGGAACTTCCACATTCAGATTCGATAGGTGA
CCGTCGACTCCAAGTGCTTTGCACAAAACAACTCCGGCCTCCCGGCCACCAGTCACACGACTCACG
GCACTACCACCCCTGACTCCCTGAGGCGGACCTGCCACTGTTCTGCATGCGAAGCTATCTAAAATT
CTGAAGCAAAGAAAGCACAGCACATGCTCCGGGACACGCGCCACCCGGCGGAAAAGGGCTCGGTGT
GGCGATCTCACAGCCGCATATCGCATTTCACAAGCCGCCCATCTCCACCGGCTTCACGAGGCTCAT
CGCGGCACGACCGCGCACGGAACGCACGCGGCCGACCCGCGCGCCTCGATGCGCGAGCCCATCCGC
CGCGTCCTCCCTTTGCCTTTGCCGCTATCCTCTCGGTCGTATCCCGTTTCTCTGTCTTTTGCTCCC
CGGCGCGCGCCAGTTCGGAGTACCAGCGAAACCCGGACACCTGGTACACCTCCGCCGGCCACAACG
CGTGTCCCCCTACGTGGCCGCGCAGCACATGCCCATGCGCGACACGTGCACCTCCTCATCCAAACT
CTCAAGTCTCAACGGTCCTATAAATGCACGGATAGCCTCAAGCTGCTCGTCACAAGGCAAGAGGCA
AGAGGCAAGAGCATCCGTATTAACCAGCCTTTTGAGACTTGAGAGTGTGTGTGACTCGATCCAGCG
TAGTTTCAGTTCGTGTGTTGGTGAGTGATTCCAGCCAAGTTTGCG

SEQ ID NO: 478, proteina - domínio sbp de consenso de polipeptídeos SPL11

HCQVEGCNLDLSSAKDYHRKHRVCEXHSKXPKVIVAGLERRFCQQCSRXXXXXXXXXXXXXXXXXX
XXXFHGLSEFDEKKRSCRRRLSDHNARRRKPQPE

FIGURA 32 CONT.

REIVINDICAÇÕES

1. Método para intensificar traços relacionados com o rendimento em plantas com relação às plantas de controle, caracterizado pelo fato de que compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo selecionado de

a. um polipeptídeo LBD, em que o dito polipeptídeo LBD compreende um domínio DUF206; ou

b. um polipeptídeo JMJC, em que dito polipeptídeo JMJC compreende um domínio JmjC; ou

c. uma caseína quinase I (CKI) em que dito CKI é selecionado de SEQ ID NO: 174 ou um ortólogo ou parálogo do mesmos; ou

d. um polipeptídeo de prolongamento-homeodomínio de homeodomínio de planta (PHDf-HD), cujo polipeptídeo de PHDf-HD compreende: (i) um domínio tendo pelo menos 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais identidades de sequência de aminoácido a um domínio zíper de leucina/prolongamento de homeodomínio de planta (ZIP/PHDf) como representado pela SEQ ID N°: 233; e (ii) um domínio tendo pelo menos 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais identidades de sequência de aminoácido a um homeodomínio (HD) como representado pela SEQ ID N°: 234; ou

e. um polipeptídeo semelhante à bHLH11, em que dito polipeptídeo semelhante à bHLH11 compreende um domínio hélice-laço-hélice; ou

f. um ASR, em que dito ASR é representado por SEQ ID NO: 397 ou um ortólogo ou parálogo do mesmo; ou

g. um polipeptídeo SPL11, em que dito polipeptídeo SPL11 compreende um domínio SBP tendo uma ordem aumentada de preferência em pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97 % ou mais identidades

de sequências à qualquer uma dentre SEQ ID N°: 456 a SEQ ID N°: 468 e SEQ ID N°: 478;

e opcionalmente selecionar para plantas tendo traços relacionados com o rendimento.

2. Método de acordo com a reivindicação 1, caracterizado pelo fato de que

a. dito polipeptídeo LBD compreende um ou mais motivos representados pela SEQ ID NO: 5, SEQ ID NO: 6 e SEQ ID NO: 7; ou

b. dito polipeptídeo JMJC compreende um motivo tendo em ordem crescente de preferência de pelo menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ou mais de identidade de seqüência à qualquer uma dentre SEQ ID N°: 79, SEQ ID N°: 80, SEQ ID N°: 81, SEQ ID N°: 82, e/ou dito domínio JmjC é representado por uma seqüência tendo em ordem crescente de preferência de pelo menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ou mais de identidade de seqüência à SEQ ID N°: 78; ou

c. dito polipeptídeo PHDf-HD compreende (i) um domínio PHD como representado por PFAM00628 e (ii) um HD como representado pela PFAM00046; ou

d. dito polipeptídeo semelhante à bHLH11 compreende um ou mais dos motivos representados por SEQ ID N°: 246, SEQ ID N°: 247, SEQ ID N°: 248, SEQ ID N°: 249, SEQ ID N°: 250, SEQ ID N°: 251, SEQ ID N°: 252; ou

e. dito polipeptídeo SPL11 além do domínio SBP compreende qualquer um ou mais dos seguintes motivos conservados:

(i) Motivo 1 como representado pela SEQ ID N°: 469 em que qualquer substituição de aminoácido conservativa e/ou 1 ou 2 na substituição conservativa são permitidos;

(ii) Motivo 2 como representado pela SEQ ID N°: 470 em que

qualquer mudança é permitida, contanto que pelo menos 4 aminoácidos tenham uma cadeia secundária polar, preferivelmente serina ou treonina e contanto que o domínio esteja localizado na extremidade do terminal N do domínio SBP;

(iii) Motivo 3 como representado pela SEQ ID: 471 em que 1 ou 2 más combinações são permitidas;

(iv) Motivo 4 como representado pela SEQ ID: 472 em que 1, 2 ou 3 más combinações são permitidas.

3. Método de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que

a. dito um dos domínios JmjC compreendido nos polipeptídeos JMJC é representado pela SEQ ID N°: 84; SEQ ID N°: 86; SEQ ID N°: 96; SEQ ID N°: 98; SEQ ID N°: 104; SEQ ID N°: 108; SEQ ID N°: 110; SEQ ID N°: 112; SEQ ID N°: 114; SEQ ID N°: 116; SEQ ID N°: 118; SEQ ID N°: 120; SEQ ID N°: 122; SEQ ID N°: 124; SEQ ID N°: 128; SEQ ID N°: 130; SEQ ID N°: 132 e SEQ ID N°: 134, cujas coordenadas de aminoácido são dadas na Tabela B4; ou

b. dito polipeptídeo PHDf-HD, quando usado na construção de uma árvore filogenética HD, tal como o descrito na Figura 13, os agrupamentos com o grupo PHDf-HD de polipeptídeos que compreende a sequência de polipeptídeo como representado pela SEQ ID N°: 180, em vez de qualquer outro grupo HD; ou

c. dito polipeptídeo PHDf-HD tem, na ordem crescente de preferência pelo menos 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais identidade de sequência ao polipeptídeo PHDf-HD como representado pela SEQ ID N°: 180 ou a qualquer uma das sequências de polipeptídeos dadas na Tabela D1 neste.

4. Método de acordo com qualquer uma das reivindicações 1 a 3, caracterizado pelo fato de que a dita expressão modulada é realizada pela

introdução e expressão em uma planta, de um ácido nucleico.

5. Método de acordo com qualquer uma das reivindicações 1 a 4, caracterizado pelo fato de que a dita expressão modulada é expressão modulada aumentada.

6. Método de acordo com qualquer uma das reivindicações 1 a 5, caracterizado pelo fato de que:

a. o dito ácido nucleico que codifica um polipeptídeo LBD codifica qualquer uma das proteínas listadas na Tabela A1 ou é uma porção de um tal ácido nucleico ou um ácido nucleico capaz de hibridizar com um tal ácido nucleico; ou

dito ácido nucleico que codifica um polipeptídeo JMJC codifica qualquer uma das proteínas listadas na Tabela B1 ou é uma porção de um tal ácido nucleico ou um ácido nucleico capaz de hibridizar com um tal ácido nucleico; ou

b. dito ácido nucleico que codifica o dito polipeptídeo CKI é uma porção da SEQ ID N°: 173 ou um ácido nucleico capaz de hibridizar com um tal ácido nucleico; ou

c. dita sequência de ácido nucleico que codifica um polipeptídeo PHDf-HD é representado por qualquer uma das sequências de ácido nucleico SEQ ID N° dadas na Tabela D1 ou uma porção destes ou uma sequência capaz de hibridização com qualquer uma das sequências de ácido nucleico SEQ ID N° dadas na Tabela D1; ou

d. dito ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 codifica qualquer uma das proteínas listadas na Tabela E1 ou é uma porção de um tal ácido nucleico ou um ácido nucleico capaz de hibridizar com um tal ácido nucleico; ou

e. dito ácido nucleico que codifica o dito polipeptídeo ASR é uma porção da SEQ ID N°: 396 ou um ácido nucleico capaz de hibridizar com um tal ácido nucleico; ou

f. dito ácido nucleico que codifica um polipeptídeo SPL11 codifica qualquer uma das proteínas listadas na Tabela G1 ou é uma porção da um tal ácido nucleico, ou um ácido nucleico capaz de hibridizar com um tal ácido nucleico.

7. Método de acordo com qualquer uma das reivindicações 1 a 6, caracterizado pelo fato de que:

a. dita sequência de ácido nucleico codificando um polipeptídeo LBD codifica um ortólogo ou parálogo de qualquer uma das proteínas dados na Tabela A1; ou

b. a dita sequência de ácido nucleico codificando um polipeptídeo JMJC codifica um ortólogo ou parálogo de qualquer uma das proteínas dadas na Tabela B1; ou

c. dita sequência de ácido nucleico codificando um polipeptídeo CKI codifica a SEQ ID N°: 174;

d. dita sequência de ácido nucleico codificando um polipeptídeo PHDf-HD codifica um ortólogo ou parálogo de acordo com qualquer uma das SEQ ID N° dadas na Tabela D1;

e. dita sequência de ácido nucleico codificando um polipeptídeo semelhante a bHLH11 codifica um ortólogo ou parálogo de qualquer uma das proteínas dadas na Tabela E1;

f. dita seqüência de ácido nucleico codificando um polipeptídeo ASR codifica SEQ ID N°: 397 ou um ortólogo ou parálogo destes;

g. dita sequência de ácido nucleico codificando um polipeptídeo SPL11 codifica um ortólogo ou parálogo de qualquer uma das proteínas dadas na Tabela G1.

8. Método de acordo com a reivindicação 7, caracterizado pelo fato de que o dito ácido nucleico codificando um polipeptídeo JMJC codifica uma SEQ ID N°: 74.

9. Método de acordo com a reivindicação 7, caracterizado pelo fato de que o dito ácido nucleico codificando um polipeptídeo SPL11 codifica uma SEQ ID N°: 428.

10. Método de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo fato de que os ditos traços intensificados relacionados com o rendimento compreendem vigor precoce aumentado e/ou rendimento aumentado, preferivelmente biomassa aumentada e/ou rendimento de semente aumentado com relação às plantas de controle, em que dito rendimento de semente aumentado compreende um ou mais de número aumentado de panícula primária; peso de semente total aumentado; número de sementes (enchidas); números de sementes ou flósculos por panícula, peso de semente milhar, taxa de enchimento de semente e/ou índice de colheita.

11. Método de acordo com qualquer uma das reivindicações de 1 a 10, caracterizado pelo fato de que os ditos traços intensificados relacionados com o rendimento são obtidos sob condições sem estresse.

12. Método de acordo com qualquer uma das reivindicações de 1 a 10, caracterizado pelo fato de que os ditos traços intensificados relacionados com o rendimento são obtidos sob condições de estresse abiótico.

13. Método de acordo com a reivindicação 12, caracterizado pelo fato de que os ditos traços intensificados relacionados com o rendimento são obtidos sob condições de deficiência de nitrogênio.

14. Método de acordo com a reivindicação 12, caracterizado pelo fato de que os ditos traços intensificados relacionados com o rendimento são obtidos sob condições de crescimento sob estresse por secura (suave).

15. Método de acordo com a reivindicação 12, caracterizado pelo fato de que os ditos traços intensificados relacionados com o rendimento são obtidos sob condições de estresse por sal.

16. Método de acordo com qualquer uma das reivindicações 4

a 15, caracterizado pelo fato de que dito ácido nucleico codificando um polipeptídeo LBD, um polipeptídeo JMJC, um polipeptídeo PHDf-HD, um polipeptídeo semelhante a bHLH11, um polipeptídeo ASR ou um polipeptídeo SPL11 está operacionalmente ligado a um promotor constitutivo, preferivelmente a um promotor GOS2, mais preferivelmente a um promotor GOS2 de arroz.

17. Método de acordo com qualquer uma das reivindicações de 4 a 15, caracterizado pelo fato de que dito ácido nucleico codificando um polipeptídeo CKI ou um polipeptídeo SPL11 está operacionalmente ligado a um promotor específico de semente, preferivelmente a um promotor WSI18, mais preferivelmente a um promotor WSI18 de arroz.

18. Método de acordo com qualquer uma das reivindicações de 1 a 16, caracterizado pelo fato de que dito ácido nucleico que codifica um polipeptídeo LBD, um polipeptídeo JMJC é de origem vegetal, preferivelmente de uma planta dicotiledônea, ainda preferivelmente da família Brassicaceae, mais preferivelmente do gênero Arabidopsis, mais preferivelmente de *Arabidopsis thaliana*.

19. Método de acordo com qualquer uma das reivindicações de 1 a 16, caracterizado pelo fato de que dito ácido nucleico que codifica um polipeptídeo PHDf-HD ou um polipeptídeo ASR é de origem vegetal, preferivelmente de uma planta monocotiledônea, ainda preferivelmente da família Poacae mais preferivelmente do gênero Oryza, mais preferivelmente de *Oryza sativa.*

20. Método de acordo com qualquer uma das reivindicações de 1 a 16, caracterizado pelo fato de que dito ácido nucleico que codifica um polipeptídeo semelhante à bHLH11 é de origem vegetal, preferivelmente de uma planta monocotiledônea, ainda preferivelmente da família Poaceae, mais preferivelmente do gênero *Triticum,* mais preferivelmente de *Triticum aestivum.*

21. Método de acordo com qualquer uma das reivindicações de 1 a 15 e reivindicação 17, caracterizado pelo fato de que dito ácido nucleico que codifica um polipeptídeo CKI é de origem vegetal, preferivelmente de uma planta dicotiledônea, ainda preferivelmente da família Solanaceae, mais preferivelmente do gênero *Nicotiana,* mais preferivelmente de *Nicotiana tabacum.*

22. Método de acordo com qualquer uma das reivindicações de 1 a 17, caracterizado pelo fato de que dito ácido nucleico que codifica um polipeptídeo SPL11 é de origem vegetal, preferivelmente de uma planta dicotiledônea, ainda preferivelmente da família Brassicaceae, mais preferivelmente do gênero Arabidopsis, mais preferivelmente de *Arabidopsis thaliana.*

23. Molécula de ácido nucleico isolada caracterizada pelo fato de que compreende qualquer uma das seguintes características:

(i) um ácido nucleico representado pela SEQ ID N°: 69, 169, 242, 401, 403, 405, 407, 409, 411, 413, 415, 417 e 448;

(ii) um ácido nucleico ou fragmento do mesmo que é complementar a qualquer uma das SEQ ID N° dadas em (i);

(iii) um ácido nucleico que codifica um polipeptídeo LBD tendo, na ordem crescente de preferência, pelo menos 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência à SEQ ID N°: 70;

(iv) um ácido nucleico que codifica um polipeptídeo SPL11 tendo, na ordem crescente de preferência, pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência à sequência de aminoácido representada pela SEQ ID N°: 449, e tendo na ordem crescente de preferência pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência a SEQ ID N°: 465;

(v) um ácido nucleico que codifica um polipeptídeo ASR

tendo, na ordem crescente de preferência, pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência à sequência de aminoácido representada por qualquer um de SEQ ID N°: 402, 404, 406, 408, 410, 412, 414, 416 e 418;

(vi) um ácido nucleico que codifica um polipeptídeo JMJC tendo, na ordem crescente de preferência, pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência à sequência de aminoácido dada na SEQ ID N°: 243;

(vii) um ácido nucleico capaz de hibridizar sob condições estringentes a qualquer um dos ácidos nucleicos dados em (i) a (vi) acima.

24. Polipeptídeo isolado, caracterizado pelo fato de que compreende:

(i) uma sequência de aminoácido representada por SEQ ID N°: 70, 170, 243, 402, 404, 406, 408, 410, 412, 414, 416, 418 e 449;

(ii) uma sequência de aminoácido tendo, na ordem crescente de preferência, pelo menos 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou 100 % de identidade de sequência à sequência de aminoácido dada na SEQ ID N°: 70, 170, 243, 402, 404, 406, 408, 410, 412, 414, 416 e 418;

(iii) uma sequência de aminoácido tendo, na ordem crescente de preferência, pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência à sequência de aminoácido representada pela SEQ ID N°: 449, e tendo na ordem crescente de preferência pelo menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais identidade de sequência a SEQ ID N°: 465;

(iv) derivados de qualquer uma das sequências de aminoácido dadas em (i).

25. Método para a produção de uma planta transgênica tendo traços relacionados com o rendimento intensificados com relação às plantas

de controle, caracterizado pelo fato de que compreende:

(i) introduzir e expressar em uma planta, um ácido nucleico que codifica um polipeptídeo LBD como definido na reivindicação 1, 2, 3, 6, 7, 8, 9 ou 24; e

(ii) cultivar a célula de planta sob condições que promovam crescimento e desenvolvimento de planta.

26. Planta ou parte da mesma, incluindo sementes, caracterizada pelo fato de ser obtenível por um método como definido em qualquer uma das reivindicações 1 a 22 e reivindicação 25, em que a dita planta ou parte da mesma compreende um ácido nucleico recombinante que codifica um polipeptídeo como definido em qualquer uma das reivindicações 1, 2, 3, 6, 7, 8, 9 ou 24.

27. Construção, caracterizada pelo fato de que compreende:

(i) ácido nucleico que codifica um polipeptídeo como definido em qualquer uma das reivindicações 1, 2, 3, 6, 7, 8, 9 ou 24;

(ii) uma ou mais sequências de controle capazes de conduzir a expressão da sequência de ácido nucleico de (i) e opcionalmente

(iii) uma sequência de terminação de transcrição.

28. Construção de acordo com a reivindicação 27, caracterizada pelo fato de que uma das ditas sequências de controle é um promotor constitutivo, preferivelmente um promotor GOS2, mais preferivelmente um promotor GOS2 de arroz.

29. Construção de acordo com a reivindicação 27, caracterizada pelo fato de que uma das ditas sequências de controle é um promotor específico de semente, preferivelmente um promotor WSI18, mais preferivelmente um promotor WSI18 de arroz.

30. Planta ou parte da mesma de acordo com a reivindicação 26, caracterizado pelo fato de ser uma planta transgênica tendo traços relacionados com o rendimento intensificados com relação às plantas de

controle, resultando de expressão aumentada de um ácido nucleico codificando um polipeptídeo como definido em qualquer uma das reivindicações 1, 2, 3, 6, 7, 8, 9 ou 24.

31. Planta ou parte da mesma de acordo com a reivindicação 26 ou reivindicação 30, caracterizado pelo fato de que dita planta é uma planta de cultivo ou uma monocotiledônea ou um cereal, tal como arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

32. Uso de um polipeptídeo como definido em qualquer uma das reivindicações 1, 2, 3, 6, 7, 8, 9 ou 24 ou de um ácido nucleico codificando tal polipeptídeo, caracterizado pelo fato de ser para intensificar traços relacionados com o rendimento com relação às plantas de controle, em particular para intensificar rendimento e/ou vigor precoce, em eu dito rendimento compreende rendimento de semente e/ou biomassa.

RESUMO

"MÉTODO PARA INTENSIFICAR TRAÇOS RELACIONADOS COM O RENDIMENTO EM PLANTAS COM RELAÇÃO ÀS PLANTAS DE CONTROLE, MOLÉCULA DE ÁCIDO NUCLEICO ISOLADA, POLIPEPTÍDEO ISOLADO, MÉTODO PARA A PRODUÇÃO DE UMA PLANTA TRANSGÊNICA, PLANTA OU PARTE DA MESMA, CONSTRUÇÃO, E, USO DE UM POLIPEPTÍDEO OU DE UM ÁCIDO NUCLEICO"

A presente invenção diz respeito a, no geral, ao campo da biologia molecualr e diz respeito a um método para intensificar traços relacionados com o rendimento e/ou melhorar várias características do desenvolvimento da planta modulando-se a expressão em uma planta de um ácido nucleico que codifica um GRP (Proteína Reguladora do Desenvolvimento). O GRP é selecionado de um domínio LOB que compreende proteína (LOB: Fronteiras de Órgão Lateral), abreviadas aqui como polipeptídeo LBD, um polipetídeo JMJC (JUMONJI-C), um polipeptídeo CKI (Caseína quinase I), uma proteína semelhante à bHLH11 (Hélice-Laço-Hélice básico 11), um polipeptídeo de prolongamento-homeodomínio de homeodomínio de planta (PHDf-HD), um polipeptídeo ASR (abscísico induzido por ácido, tensão e maturação) e/ou um polipeptídeo de fator de transcrição semelhante à proteína de ligação do promotor Squamosa 11 (SPL11). A presente invenção também diz respeito à plantas tendo expressão modulada de um ácido nucleico que codifica um GRP, cujas plantas tem características de desenvolvimento melhoradas com relação às plantas do tipo selvagem ou outras plantas de controle. A invenção também fornece novos ácidos nucleicos GRP e polipeptídeos GRP bem como construções úteis nos métodos da invenção.